



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> C12N 15/12, 15/62, 5/10, 1/21, C07K 14/47, 16/18, C12Q 1/68, A61K 31/70, 38/17, 39/395, 48/00, G01N 33/577	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/53071</b> <b>(43) International Publication Date:</b> 26 November 1998 (26.11.98)
<b>(21) International Application Number:</b> PCT/US98/10547 <b>(22) International Filing Date:</b> 22 May 1998 (22.05.98)  <b>(30) Priority Data:</b> 60/047,490 23 May 1997 (23.05.97) US 60/047,491 23 May 1997 (23.05.97) US  <b>(71) Applicant (for all designated States except US):</b> BIOGEN, INC. [US/US]; 14 Cambridge Center, Cambridge, MA 02142 (US).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> SANICOLA-NADEL, Michele [US/US]; 4 Maple Road, Winchester, MA 01890 (US). HESSION, Catherine, A. [US/US]; 35 Otis Hill Road, Hingham, MA 02043 (US). WEI, Henry [CN/US]; 62 Farwell Street, Newton, MA 02160 (US). CATE, Richard, L. [US/US]; 64 Arrowhead Road, Weston, MA 02193 (US).  <b>(74) Agent:</b> FENTON, Gillian, M.; Biogen, Inc., 14 Cambridge Center, Cambridge, MA 02142 (US).		<b>(81) Designated States:</b> AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> MODULATORS OF TISSUE REGENERATION  <b>(57) Abstract</b>  Proteins which are upregulated in injured or regenerating tissues, as well as the DNA encoding these proteins, are disclosed, as well as therapeutic compositions and methods of treatment encompassing these compounds.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## MODULATORS OF TISSUE REGENERATION

### Related Applications

This is a continuation-in-part of prior U.S. Provisionals S.N. 60/047,490 and S.N. 60/047,491, both filed May 23, 1997. The teachings of both earlier-filed Provisional patent applications are incorporated herein by reference.

### Field of the Invention

The invention relates generally to gene expression products, including nucleic acids (e.g., RNAs) and polypeptides that are upregulated in injured or regenerating tissues. Further, the invention relates generally to cDNAs and other nucleic acids encoding polypeptides that are upregulated in injured or regenerating tissues.

### Background of the Invention

The processes of tissue development during embryogenesis, and of tissue repair and/or regeneration following injury or insult are presently the topic of intense investigation. Both processes involve a dynamic remodeling of tissue architecture, which is triggered and mediated by numerous biological interactions, including cell-cell contact, cell-matrix contact, release of soluble biological response modifiers, synthesis of structural components, and many other changes in cell phenotype, including changes in gene expression. Many of the factors involved in tissue genesis and in the response to tissue insult remain unknown or poorly understood.

Several systems have been developed for modeling the events triggered by insult to particular tissues, and/or by particular types of insult. For example, several investigators have described events occurring in mammalian kidney tissue exposed to an ischemia-reperfusion insult. The kidney is able to repair damage to the proximal tubule epithelium through a complex series of events involving cell death, proliferation of surviving proximal tubule epithelial cells, formation of poorly differentiated regenerative epithelium over the

denuded basement membrane, and differentiation of the regenerative epithelium to form fully functional proximal tubule epithelial cells (Wallin et al., Lab. Invest. 66:474-484, 1992; Witzgall et al., Mol. Cell. Biol. 13:1933-1942, 1994; Ichimura et al., Am. J. Physiol. 269:F653-662, 1995; Thadhani et al., N. Engl. J. Med. 334:1448-1460, 1996). Growth factors such as IGF, EGF, and HGF have been implicated in this process of repair, as has the endothelial cell adhesion molecule ICAM-1. However, the mechanisms by which the tubular epithelial cells are restored, either functionally and morphologically, remain poorly understood.

There accordingly is a need for an improved understanding of the biological processes of tissue repair or regeneration. That is, there is a need to identify factors that are indicative of the initiation, development and resolution of normal wound healing or normal biological response to tissue insult. Similarly, there is a need to identify factors that are indicative of pathologies of abnormal responses to tissue insult or other stimulus, including autoimmune or other dysregulation pathologies, as well as pathologies associated with abnormal tissue growth including neoplastic growth. Further, there is a need to identify factors that constitute appropriate targets for therapeutic intervention to stimulate, modulate, enhance, suppress or otherwise manipulate biological responses to tissue insult. Similarly, there is a need to identify factors that constitute appropriate targets for therapeutic intervention to manipulate abnormal or dysregulated responses to tissue insult or other stimulus, including abnormal tissue growth associated with cancer (neoplasia) and, conversely, abnormal tissue quiescence associated with degenerative diseases.

### **Summary of the Invention**

It is an object of this invention to identify factors that are indicative of the initiation, development and resolution of normal wound healing or normal biological response to tissue insult. A specific object of this invention is to identify products (e.g., RNAs, cDNAs, polypeptides) of genes that are upregulated in the context of healing tissue, or in tissue exposed to insult or injury. A particular object is to identify products of genes that are upregulated in mammalian kidney tissue exposed to insult or injury, such as ischemia-



reperfusion injury. A second specific object of this invention is to identify products of genes that are upregulated in pathological conditions associated with abnormal responses to tissue insult or other stimulus, including autoimmune or other dysregulation pathologies, as well as pathologies associated with abnormal tissue growth including neoplastic growth.

- 5 A third specific object is to identify gene products that constitute appropriate targets for therapeutic intervention to stimulate, modulate, enhance, suppress or otherwise manipulate biological responses to tissue insult. A particular object is to identify gene products for therapeutic intervention to manipulate responses of mammalian kidney tissue to tissue injury, such as ischemia-reperfusion injury. A fourth specific object is to identify gene
- 10 products that constitute appropriate targets for therapeutic intervention to manipulate abnormal or dysregulated responses to tissue insult or other stimulus, including abnormal tissue growth associated with cancer (neoplasia) and, conversely, abnormal tissue quiescence associated with degenerative diseases. Thus, other particular objects of the invention include the identification of gene products as targets for therapeutic intervention
- 15 in the clinical management (including prophylaxis, maintenance and treatment) of kidney diseases, including diseases involving renal failure, and of cancers derived from or affecting renal tissue.

The present invention rests on the discovery that the expression levels of numerous genes, including many genes heretofore unknown and/or uncharacterized, are upregulated

20 in mammalian kidney tissue that has been exposed to tissue insult or injury. More specifically, the present invention rests on the discovery of a novel class of Kidney Injury-associated Molecules (each member of the class is henceforth called a "KIM"). KIMs are polypeptides encoded by genes whose expression levels are upregulated in the mammalian kidney upon exposure to tissue injury, particularly after exposure to ischemia-reperfusion

25 injury. Any KIM is useful as an indicator of tissue status (preferably renal tissue status) or of a change therein (e.g, exposure to tissue injury, or the occurrence or stage of a tissue repair process). Any KIM further is useful as a target for therapeutic intervention to manipulate, whether by an agonist or an antagonist, a normal or abnormal tissue response to insult or other stimulus.

The invention accordingly provides, in a first aspect, purified and isolated nucleic acid molecules encoding all or a unique fragment of a KIM. In one embodiment, the present nucleic acids are RNAs. In another, they are DNA molecules, such as cDNAs. In another embodiment, the invention provides the complementary strands of nucleic acids encoding all or a unique fragment of a KIM. In another embodiment, the invention provides nucleic acids (preferably DNAs) that hybridize under low or, preferably, high stringency conditions to any of the foregoing nucleic acids. In still another embodiment, the invention provides nucleic acids (preferably DNAs) which, but for the degeneracy of the genetic code, would hybridize to any of the foregoing nucleic acids. In some embodiments, a nucleic acid encoding all or a unique fragment of a KIM is an engineered (recombinant) nucleic acid, optionally in operative association with an expression control element or other regulatory element. In other embodiments, a nucleic acid encoding all or a unique fragment of a KIM is an antisense nucleic acid sufficient, when internalized within a cell, to disrupt expression of a cellular KIM gene.

Specific novel KIMs (and unique fragments thereof) of the present invention are coded for by nucleic acids having the sequences disclosed herein in TABLE 1 and in the Sequence Listing. Other specific KIMs (and unique fragments thereof) are coded for by nucleic acids having sequences that are degenerate variants of any of the KIM sequences set forth in TABLE 1 and in the Sequence Listing. Still other specific KIMs (and fragments) are coded for by nucleic acids that are substantially similar to (homologous to) any of the KIM sequences in TABLE 1 and in the Sequence Listing. Such KIMs are defined herein as variants of the disclosed novel KIM sequences. In some embodiments, the present nucleic acid encodes a chimeric polypeptide comprising a novel KIM-encoding sequence (i.e., a sequence encoding part or all of a KIM) fused to a non-KIM sequence. Thus, the invention provides nucleic acids encoding novel KIM fusion proteins, non-limiting examples of which include KIM polypeptides fused to a secretable leader polypeptide, an immunoglobulin polypeptide, a binding pair partner (e.g., avidin, GST), a toxin or toxoid (e.g., ricin, tetanus), an enzyme (preferably one for which a detectable substrate is available, e.g., alkaline phosphatase, horseradish peroxidase, luciferase), or

other detectable polypeptide (e.g., green fluorescent protein). Additional specific KIMs (and unique fragments thereof) are coded for by nucleic acids referred to in TABLE 2. It is believed that, in most to all instances, these specific molecules have not heretofore been appreciated to be KIMs.

5 In a second aspect, the invention provides a vector having a KIM-encoding nucleic acid inserted therein. In some embodiments, the vector is a biologically functional plasmid or viral DNA vector. In other embodiments, the vector is a retroviral vector.

10 In a third aspect, the invention provides a prokaryotic or eukaryotic host cell comprising an internalized vector having a KIM-encoding nucleic acid insert. The present host cell provides intracellular means for producing (synthesizing, folding, processing, or secreting) a KIM polypeptide of the present invention.

15 In a fourth aspect, the invention provides a process for the production of a KIM polypeptide. The present production process includes growing a host cell of the invention under culture conditions sufficient for the production of polypeptides of vector origin, and recovering an expressed KIM polypeptide.

20 In a fifth aspect, the invention provides a purified and isolated novel KIM polypeptide, preferably substantially free of non-KIM polypeptides or proteins. In some embodiments, the present polypeptide is a full-length polypeptide, i.e., a polypeptide corresponding to the full-length open reading frame of a novel KIM-encoding cDNA. In other embodiments, the present polypeptide is a unique fragment of the full-length polypeptide. Thus, specific novel KIM polypeptides (and unique fragments thereof) of the present invention are coded for by nucleic acids having the sequences disclosed herein in TABLE 1 and in the Sequence Listing. Other specific novel KIM polypeptides of the present invention are variants of the disclosed novel KIM sequences, including without  
25 limitation splice variants, truncation variants, and substitution variants. In some embodiments, the invention provides a chimeric polypeptide comprising a novel KIM polypeptide fused (preferably via a peptide bond) to a non-KIM polypeptide. Thus, the invention provides novel KIM fusion proteins, exemplified by the above-mentioned fusion

constructs. Of course, the invention also provides conjugated or derivatized novel KIM polypeptides, including without limitation detectable conjugates, imageable conjugates, radiolabeled conjugates, and toxin conjugates. In still other embodiments, the invention provides chimeric polypeptides, fusion proteins and conjugates comprising a polypeptide  
5 appreciated herein as being a KIM (see TABLE 2).

In a sixth aspect, the invention provides an antibody that binds selectively to a KIM polypeptide. Preferably, the antibody is a monoclonal antibody (or an engineered derivative thereof) produced by conventional means from a hybridoma derived from splenocytes of an animal immunized with a KIM polypeptide of the present invention. In  
10 some embodiments, the present antibody is conjugated or derivatized with a detectable moiety, toxin, imageable compound or radionuclide.

In a seventh aspect, the invention provides a pharmaceutical composition comprising a therapeutically effective amount of a KIM nucleic acid (e.g., an antisense nucleic acid); a KIM vector; a KIM polypeptide; a KIM fusion protein; a KIM-binding  
15 antibody (also referred to as an anti-KIM antibody), dispersed, dissolved or otherwise suspended in a physiologically acceptable carrier, vehicle, solvent or excipient.

In an eighth aspect, the invention provides methods and kits for detecting, either qualitatively or quantitatively, KIM expression and/or KIM polypeptide. As disclosed herein, KIMs are upregulated by exposure of mammalian kidney (renal) tissue to insult or  
20 injury, particularly ischemia-reperfusion injury. Thus, any KIM (or combination or panel thereof) can be used as an indicator of tissue status (preferably renal tissue status), or of a change therein, including without limitation exposure to tissue injury, the occurrence or stage of an injury process, the occurrence or stage of a tissue repair or regeneration process, or the occurrence of an abnormal tissue response to injury or other stimulus, such as an  
25 autoimmune response or an abnormal proliferative response (e.g., a neoplastic response).

In one embodiment, the invention provides a detection method and kit for diagnosis, prognosis, staging, or monitoring of renal injury or of renal disease or of the effectiveness of therapy therefor. Some specific embodiments involve detecting and/or

measuring the concentration of one or more KIM polypeptides in serum, urine, or urine sediment of an individual (a mammal, preferably a human) afflicted with or at risk of developing renal injury or an impairment of renal function. Some such embodiments involve the use of at least one KIM polypeptide, anti-KIM antibody or conjugate thereof as elements of a reagent kit for immunoassay according to standard techniques. For present purposes, any KIM disclosed herein or identified as such herein can be used as a kit element, along with other elements for KIM detection. In other instances, such as where the detected KIM has biological or enzymatic activity, the kit can include reagents for detecting KIM activity, e.g., by enzyme assay. Other specific embodiments involve detecting and/or measuring the level of expression of one or more KIM-encoding nucleic acids in renal cells in a kidney biopsy, or in cells shed into urine or urine sediment of an individual afflicted with or at risk of developing renal injury or an impairment of renal function. Some such embodiments involve the use of a nucleic acid (e.g., RNA or DNA) encoding a KIM polypeptide, or a unique fragment thereof, or a probe nucleic acid capable of hybridizing to nucleic acid encoding a KIM, as elements of a reagent kit for hybridization assay according to standard techniques. As above, any KIM nucleic acid disclosed herein or appreciated herein as such can be used as a hybridization kit element, along with other reagents for detection of hybridized nucleic acids.

In another embodiment, the invention provides a method and kit for diagnosis, prognosis, staging or monitoring of an abnormal response of renal tissue to tissue injury or other stimulus, including an autoimmune response or an abnormal proliferative response, such as neoplasia arising from or affecting renal tissue. Some specific embodiments involve detecting and/or measuring the concentration of one or more KIM polypeptides in serum, urine, or urine sediment of an individual (a mammal, preferably a human) afflicted with or at risk of developing renal injury or an impairment of renal function. Some such embodiments involve the use of at least one KIM polypeptide, anti-KIM antibody or conjugate thereof as elements of a reagent kit for immunoassay according to standard techniques. For present purposes, any KIM disclosed herein or identified as such herein can be used as a kit element, along with other elements for KIM detection. In other instances,

such as where the detected KIM has biological or enzymatic activity, the kit can include reagents for detecting KIM activity, e.g., by enzyme assay. Other specific embodiments involve detecting and/or measuring the level of expression of one or more KIM-encoding nucleic acids in cells present in a kidney biopsy, or in cells shed into urine or urine

5 sediment of an individual afflicted with or at risk of developing renal injury or an impairment of renal function. Some such embodiments involve the use of a nucleic acid (e.g., RNA or DNA) encoding a KIM polypeptide, or a unique fragment thereof, or a probe nucleic acid capable of hybridizing to nucleic acid encoding a KIM, as elements of a reagent kit for hybridization assay according to standard techniques. As above, any KIM  
10 nucleic acid disclosed herein or appreciated herein as such can be used as a hybridization kit element, along with other reagents for detection of hybridized nucleic acids.

In a ninth aspect, the invention provides a method and reagent for imaging tissues, either *in vitro* or *in vivo*. In particular, the invention provides a method and reagent for imaging the presence, extent or severity of kidney injury or of kidney tissue repair or  
15 regeneration. Similarly, the invention provides a method and reagent for imaging the presence, extent, severity or stage of an abnormal response to tissue injury or other stimulus, such as autoimmunity or neoplasia, particularly neoplasia arising from or affecting renal tissue. The present invention also provides a method for targetting an imageable compound to cells or tissue expressing or producing a KIM. The present  
20 method involves the step of contacting cells or tissue with a detectable (e.g., imageable) KIM-binding reagent, which in some embodiments is an anti-KIM antibody or conjugate thereof, or a KIM fusion protein. In other embodiments, the detectable KIM-binding reagent is a nucleic acid of the invention (e.g., a probe or antisense nucleic acid) labeled with a radionuclide or other imageable compound. According to the present method, cells  
25 expressing or producing a KIM are visualized (imaged) by detecting the presence and/or location of an accumulation of the KIM-binding reagent. For imaging *in vivo*, the KIM-binding reagent is administered, by any appropriate route, to an individual (a mammal, preferably a human) suspected of harboring an imageable locus of KIM expression and/or production. Without being limited hereby, it is believed that the present method can be

used to detect the presence and/or location of a tissue mass (e.g., a tumor) abnormally producing or expressing a KIM.

In a tenth aspect, the invention provides methods for treating an individual (a mammal, preferably a human) afflicted with or at risk of developing a disease or condition contributed to or associated with a dysfunction or dysregulation of a KIM gene or protein. Further, the invention provides methods for treating an individual afflicted with or at risk of developing a disease or condition beneficially affected by therapeutic administration of a KIM protein. The present methods involve the step of administering to the individual a therapeutically effective amount of a KIM polypeptide, variant or fusion protein thereof, or, conversely, of an anti-KIM antibody. It is expected that such compounds will be useful in therapeutic methods which manipulate, e.g., stimulate or inhibit, biological responses that are dependent on KIM function.

In an eleventh aspect, the invention provides a method for inhibiting the growth of KIM-expressing tumor cells, involving the step of contacting the cells with an anti-KIM antibody, conjugated to a toxin or radionuclide. In an alternative embodiment, the method involves the step of contacting the cells with an antisense KIM nucleic acid that is sufficient to suppress or disrupt expression of a KIM gene in the tumor cells.

In a twelfth aspect, the invention provides a method of gene therapy. The present method involves the administration of a vector capable of directing the production of a KIM, to an individual afflicted with or at risk of a renal disorder, disease or injury. As a result of the present method, growth of new tissue, preferably renal tissue, is stimulated, or survival of existing tissue, preferably renal tissue, is promoted.

The foregoing and other objects, features, aspects and advantages of the present invention, as well as the invention itself, will be more fully understood from the following description of preferred embodiments.

### **Detailed Description of the Invention**

The discovery, reported herein, of a novel class of Kidney Injury-related Molecules (KIMs) was made by analyzing differences in mRNA expression between normal adult mammalian kidneys, and kidneys in the process of regenerating following exposure to tissue insult (specifically, to ischemia-reperfusion injury). Two established techniques were used for this purpose: representational difference analysis (RDA), and suppression subtractive hybridization (SSH). Both techniques were used to assess cDNAs isolated from various timepoints, e.g., 48 hours, after the onset of ischemia. In these studies, the normal adult kidney material was isolated from sham-operated subjects. Both techniques resulted in the depletion of cDNAs which are common to both postischemic and to normal kidney samples, leaving a pool of cDNAs which are significantly expressed only in injured or regenerating kidney tissue. This pool likely contains cDNAs corresponding to genes that encode proteins involved in the injury process, and proteins involved in tissue repair or regeneration processes. Therefore, cDNAs isolated from the pool likely encode parts or all of proteins likely to be therapeutically beneficial for treatment or prophylaxis of tissue injury, especially renal injury. Several cDNA clones have been obtained, sequenced and characterized.

### **Selected Definitions**

A "KIM protein", herein used synonymously with "KIM", is any protein or polypeptide encoded by mRNA which is selectively upregulated following injury to a kidney. One group of KIM proteins of interest includes those coded for by mRNA which is selectively upregulated at any time within one week following any insult which results in injury to renal tissue. Examples of times at which such upregulation might be identified include 10 hours, 24 hours, 48 hours or 96 hours following an insult. Examples of tissue insults include toxin exposure, hypoxia, hyperoxia, hemodynamic disruption, ischemia, reperfusion, or mechanical compression. Many different types of proteins fall within the KIM class, including cell surface proteins (e.g., ligands or counter-receptors involved in cell-cell or cell-matrix interactions), secreted proteins (e.g., diffusable biological response modifiers, such as growth factors, differentiation factors, survival factors and the like),



intracellular proteins (e.g., elements of a signalling pathway), and nuclear proteins (e.g., transcription factors).

5 A "KIM ligand" is any molecule which noncovalently and specifically binds to a KIM protein. Such a ligand can be a protein, peptide, steroid, antibody, amino acid derivative, or other type molecule, in any form, including naturally-occurring, recombinantly produced, or otherwise synthetic. A KIM ligand can be in any form, including soluble, membrane-bound, or part of a fusion construct with immunoglobulin, fatty acid, or other moieties. The KIM ligand may be an integrin. A membrane-bound KIM ligand can act as a receptor which, when bound to or associated with KIM, triggers a  
10 cellular response. In some interactions, a KIM may associate with a plurality of KIM ligands, or may associate with a KIM ligand as part of a complex with one or more other molecules or cofactors. In a situation where both the KIM and the KIM ligand are bound to cell membranes, the KIM may associate and react with KIM ligand which is bound to the same cell as the KIM, or it may associate and react with KIM ligand be bound to a  
15 second cell. Where the KIM ligation occurs between molecules bound to different cells, the two cells may be the same or different with respect to cellular type or origin, phenotypic or metabolic condition, or type or degree of cellular response (e.g., growth, differentiation or apoptosis) to a given stimulus. "KIM ligation" refers to the contact and binding of KIM with a KIM ligand.

20 A "unique fragment" of a nucleic acid means any fragment of sufficient length to have a sequence likely to be substantially unique in a mammalian genome. Thus, a unique fragment generally means an oligonucleotide at least 16 nucleotide bases in length. Similarly, a "unique fragment" of a polypeptide means any fragment of sufficient length to have an amino acid sequence likely to be substantially unique to a given mammalian  
25 protein, such as a KIM. Thus, a unique peptide fragment generally means a peptide at least 7 amino acids in length.

A "KIM variant" means a KIM whose sequence differs from a sequence disclosed herein by the presence of one or more internal or terminal insertions, deletions or substitutions of a nucleotide (when referring to KIM nucleic acids) or of an amino acid

(when referring to KIM polypeptides). Preferably, the KIM variant is "substantially similar" to the corresponding disclosed KIM, or to a unique fragment thereof. That is, the KIM variant is "homologous" to the corresponding disclosed KIM. "Substantially similar" or "homologous" variants are structurally similar to the corresponding disclosed KIM. Further, "substantially similar" or "homologous" variants have sufficient functional similarity to the corresponding disclosed KIM that they share one or more of the KIM's biological properties or functions (e.g., binding to a receptor or ligand, triggering of biological responses, transport of a metabolite, catalysis of a substrate, or the like). A KIM variant can be naturally occurring or synthesized or produced by routine techniques, such as molecular engineering techniques.

By "alignment of sequences" is meant the positioning of one sequence, either nucleotide or amino acid, with that of another, to allow a comparison of the sequence of relevant portions of one with that of the other. Generally, sequences are aligned using the GAP and BESTFIT programs, which are based on the teachings of Needleman et al. (J. Mol. Biol. 48:443-453, 1970), Smith et al. (Adv. Appl. Math. 2:482-489, 1981), and Rechid et al. (CABIOS 5:107-113, 1989). Generally, when a homologous variant of a KIM is aligned with the corresponding KIM polypeptide, it will share amino acids that contribute to the KIM's three dimensional structure, such as cysteine residues. Although the relative positions of cysteine residues is generally conserved in the variant sequence, homologous or functionally equivalent sequences can include functionally equivalent arrangements of the cysteines, including arrangements comprising amino acid insertions or deletions which alter the linear arrangement of the cysteines, but do not materially impair their relationship or ability to form disulfide bonds in the folded structure of the KIM protein. Therefore, minor internal gaps and amino acid insertions, such as occur in splice variants, are ignored when aligning sequences herein.

"Sequence homology", "percent (%) homology", "sequence similarity" and "percent (%) similarity" are used interchangeably herein, and refer to the *sum* of the percentage of residues (whether nucleotides or amino acids) that are, when aligned with a reference sequence, the same as the corresponding reference residues, *and* those that are,

when aligned, conservative substitutions for the corresponding reference residues.

“Sequence identity” and “percent (%) identity” also are used interchangeably herein, and refer to the percentage of residues (whether nucleotides or amino acids) that are, when aligned with a reference sequence, the same as the corresponding reference residues. Both sets of terms are used according to their definitions in Altschul et al. (1990), J. Mol. Biol. 215:403-410 and in the Basic Local Alignment Search Tool (BLAST) algorithm described therein. For present purposes, the algorithm gap weight is set at 3.0 and the length weight is set at 0.1.

“Hybridization” means the formation of a duplex nucleic acid molecule, in which nucleotide bases of a first polynucleotide strand bind noncovalently with cognate nucleotide bases of a second polynucleotide strand. Generally, such noncovalent binding occurs in DNA only between adenosine (A) and thymidine (T) bases, and guanosine (G) and cytosine (C) bases. In RNA, binding occurs only between A and uracil (U), and G and C. The apposition of non-cognate bases in a duplex nucleic acid molecule, e.g., A with C, is termed a “mismatch” pairing. Under high stringency hybridization conditions, mismatches occur rarely in a given duplex nucleic acid molecule. Low stringency conditions permit the occurrence of some mismatches. Exemplary conditions which promote DNA hybridization, termed “hybridization conditions”, include 6.0X sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0X SSC at 50°C. The salt concentration in the wash step can be selected from a low stringency wash of about 2.0X SSC at 50°C, to a high stringency wash of about 0.2X SSC at 50°C. In addition, the temperature in the wash step can be selected from a low stringency wash at room temperature, about 22°C, to a high stringency wash at about 65°C.

A “KIM agonist” is a molecule which can specifically trigger a cellular response normally triggered by the interaction of KIM with a KIM ligand. A KIM agonist can be a KIM variant, or a specific antibody to KIM, or a soluble form of the KIM ligand.

A “KIM antagonist” is a molecule which can specifically associate with a KIM ligand or a KIM, thereby blocking or otherwise inhibiting KIM binding to the KIM ligand. The antagonist binding blocks or inhibits cellular responses which would otherwise be

triggered by ligation of the KIM ligand with KIM or with a KIM agonist. Examples of KIM antagonists include certain KIM variants, KIM fusion proteins and specific antibodies to a KIM ligand or KIM.

5 "Isolated" refers to a condition in which a nucleic acid or polypeptide of the present invention is essentially free of other nucleic acids, polypeptides, or of other contaminants with which it might normally be found in nature, and as such exists in a form not found in nature.

10 "Substantially pure" refers to a condition in which a nucleic acid or polypeptide of the present invention is separated from other nucleic acids, polypeptides, or other contaminants, particularly naturally occurring contaminants, that interfere with the ability to detect, visualize or isolate the the nucleic acid or polypeptide of the invention, or that interfere with a biological function or property thereof. A substantially pure nucleic acid or polypeptide of the invention is not generally found in nature.

15 A "chemical derivative" of another molecule contains one or more additional chemical or biochemical moieties not found naturally in association with the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 16th ed., Mack  
20 Publishing Co., Easton, Penn. (1980).

Other terms used herein, including "antisense DNA", "antisense probe", "cloning", "cDNA", "cDNA library", "DNA polymorphism", "expression", "gene", "hybridoma", "plasmid", "probe", "labeled", "recombinant", "host cells", "transformed", "transfected", "vector", and the like are used in the sense of their art-recognized meanings in the fields of  
25 cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA or genetic engineering, and immunology. Such meanings are determined by consultation of one or more of the following widely available texts: Molecular Cloning, A Laboratory Manual, 2nd Ed. (Sambrook, Fritsch and Maniatis, eds.),

- Cold Spring Harbor Laboratory Press, 1989; DNA Cloning, Volumes I and II (Glover, ed.), 1985; Oligonucleotide Synthesis (Gait, ed.), 1984; U.S. Patent 4,683,195, Mullis et al., invs.; Nucleic Acid Hybridization (Hames & Higgins, eds.), 1984; Transcription and Translation (Hames & Higgins, eds.), 1984; Culture of Animal Cells (Freshney), Alan R. Liss, publ., 1987; Immobilized Cells and Enzymes (IRL Press), 1986; A Practical Guide to Molecular Cloning (Perbal), 1984; Current Protocols in Molecular Biology, Wiley & Sons, publ., 1989; Methods in Enzymology, Academic Press, New York NY (especially Volumes 154 and 155); Gene Transfer Vectors for Mammalian Cells (Miller and Calos, eds.), Cold Spring Harbor Laboratory Press, 1987; Immunochemical Methods in Cell and Molecular Biology (Mayer and Walker, eds.), Academic Press, London, 1987; Handbook of Experimental Immunology, Volumes I-IV (Weir and Blackwell, eds.), 1986; and, Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, 1986.

#### Compounds of the Invention

- Specific, novel KIMs discovered through RDA or SSH techniques are set forth below in TABLE 1 and in the Sequence Listing. TABLE 1 lists KIMs for which potential full length (PFL) cDNA clones were obtained. Also shown are the predicted polypeptide sequences encoded in the open reading frames (ORFs) of many of the newly discovered KIMs. In addition, TABLE 1 provides, where relevant, the subtracted amplified clone (SAC) designations and the PFL location(s) of partial sequences, many of which were reported in prior U.S. Provisional S.N. 60/047,490 and 60/047,491.

TABLE 1

SEQ. NO.	PFL	CLONE DESIGNATION	OPEN READING FRAME (OFR)	
			LOCATION	SEQ. ID #
# 3	HW011	HW011		
# 4	HW012	HW012	39..671	# 5
# 6	HW013	HW013		
# 7	HW014	HW014rev	175..804	# 8
# 9	HW015	HW15	11..2176	# 10
# 12	HW017	SAC_24091		
# 13	HW018 HW101	SAC_23880	86..1102	# 14
# 16	HW033	SAC_23901		
# 17	HW034	SAC_23897	91..837 3101..4162	# 18 # 19
# 20	HW035 HW112	SAC_24468		
# 21	HW036	SAC_24406		
# 22	HW037 HW102	SAC_24354	280..1422	
# 26	HW040	SAC_24520	481..2433	# 27
# 28	HW041	SAC_24317		
# 29	HW042 HW113	SAC_24017	8..889	# 30
# 31	HW043	SAC_24533	8..634	# 32
# 33	HW044	SAC_24216	24..1100	# 34
# 35	HW045	SAC_24028	9..1628	# 36
# 37	HW046	SAC_24036	9..1160	# 38
# 39	HW047	SAC_23915	8..511	# 40
# 42	HW050 HW109	SAC_24644	7..1326	# 43
# 44	HW051	SAC_24170	114..1505	# 45
# 46	HW052	SAC_24882	247..765	# 47
# 49	HW055	SAC_24449	214..1329	# 50
# 51	HW056	SAC_24326	8..523	# 52
# 53	HW057	SAC_23926	7..1023	# 54
# 55	HW059	SAC_24457	9..1271	# 56
# 57	HW061	SAC_24029	7..1422	# 58
# 61	HW069	SAC_24477	1360..1893	# 62
# 64	HW073	SAC_24456	9..1223	# 65
# 66	HW074	SAC_24464	7..300	# 67
# 68	HW075 HW117	SAC_24466		
# 69	HW076	SAC_24409		
# 71	HW080	SAC_24033		
# 72	HW082	SAC_24469	150..569	# 73
# 75	HW084	SAC_24854		
# 76	HW088	SAC_24336	9..686	# 77
# 78	HW089	SAC_24461	8..1006	# 79

TABLE 1, continued

SEQ. NO.	PFL	CLONE DESIGNATION	OPEN READING FRAME (OFR)	
			LOCATION	SEQ. ID #
# 80	HW090 HW118	SAC_24197	335 . 937	# 81
# 82	HW092	SAC_24320		
# 83	HW093	SAC_24538	8 . . 1357	# 84
# 85	HW094	SAC_23896	255 . . 1238	# 86
# 87	HW095	SAC_23802	7 . . 876	# 88
# 89	HW096	NONE	295 . . 1302	# 90

Additional KIMs, which are believed to be related to (e.g., substantially similar to) known molecules also were obtained through RDA or SSH techniques, and are set forth in

5 TABLE 2 and in the Sequence Listing. It is believed that few to none of these molecules were previously appreciated to be KIMs, i.e., to have relevance to tissue injury or repair processes, particularly in the kidney. TABLE 2 lists these KIMs along with their clone designations. Partial sequences of some of the cDNA clones listed below were also set forth in the priority document. Also shown are the name(s) and GENBANK accession

10 numbers of the known genes to which these KIMs may be related.

TABLE 2

SEQ ID #	DESIGNATION	RELATED GENE	Accession #
92	23798	Rat alpha 1A (1D) adrenergic receptor	M60654
93	23800	Rat N27	V30789
94	23801	Rat alpha-2mu globulin-related protein	X13295
95	23833	Rat contrapsin-like protease inhibitor related protein	X16359
96	23836	Rat Fit-1	V04319
97	23853	Rat tumor-associated antigen (pE4)	L12025
98	23906	Rat sulfated glycoprotein 2	X13231
99	23907	Mouse secreted protein (p85)	L33416
100	23916	Human Differentiation-dependent A4 protein	L09604

TABLE 2, continued

SEQ ID #	DESIGNATION	RELATED GENE	Accession #
101	23922	Rat manganese-containing superoxide dismutase	Y00497
102	24030	Rat heme oxygenase	J02722
103	24078	Rat tissue inhibitor of metalloproteinase 1 (TIMP1)	U16022
104	24105	Rat alpha 2- macroglobulin	J02635
105	24140	Rat Cathepsin D	X54467
106	24142	Rat lysozyme	L12458
107	24192	Mouse MAMA mRNA	X67809
108	24296	Rat transketolase	U09256
109	24445	Rat heat stable antigen CD24/ Elongation factor 1	Z11531
110	24527	Cathepsin L	S85184
111	24540	Mouse endothelial monocyte-activating polypeptide 1	U41341
112	24623	Rat spleen thymosin beta4	K01334
1	HW010	Human semaphorin (CD100)	U60800
11	HW016	Mouse ribonucleotide reductase M2 subunit	M14223
15	HW032	Human S100C	D49355
24	HW038	Human splicing factor SF3a120	X85237
25	HW039	Annexin 11	M82802
41	HW049	Mouse u22 snoRNA host gene (UHG)	U40654
48	HW054	Rat myelin oligodendrocyte protein (MOG)	M99485
59	HW062	Mouse proliferation-associated protein 1	U43918
60	HW066	Human cleavage stimulation factor 50kd subunit	L02547
63	HW070	Human BRCA2 region mRNA sequence CG037	U50523



TABLE 2, continued

SEQ ID #	DESIGNATION	RELATED GENE	Accession #
70	HW078	Human G protein gamma-10 subunit	P50151
74	HW083	Rat ins-1 winged helix	P97691
91	HW097	Human Na <sup>+</sup> channel 2	U78181

Any of the foregoing KIM cDNAs can be inserted into a vector, liposome or other carrier vehicle for internalization and production in a host cell. Furthermore, the invention encompasses derivatives and variants of each of the foregoing KIMs as listed in TABLES 1 and 2.

One embodiment of the invention provides soluble variants of a KIM protein that is usually synthesized as a membrane associated protein. Soluble variants lack at least a portion of the transmembrane or intra-membrane section of a native KIM protein. In some examples, the soluble variant lacks the entire transmembrane or intra-membrane section of a native KIM protein. Soluble variants include fusion proteins which encompass derivatives of KIM proteins that lack at least a portion of the transmembrane or intra-membrane section of a native KIM protein. All types of KIM fusion proteins are included, particularly those which incorporate his-tag, Ig-tag, and myc-tag forms of the molecule. These KIM fusions may have characteristics which are therapeutically advantageous, such as the increased half-life conferred by the Ig-tag. Also included are fusion proteins which incorporate portions of selected domains of the KIM protein.

Variants can differ from a naturally occurring KIM protein in amino acid sequence or in ways that do not involve sequence, or both. Variants in amino acid sequence are produced when one or more amino acids in naturally occurring KIM protein is substituted with a different natural amino acid, an amino acid derivative or non-natural amino acid. Particularly preferred substitution variants include naturally occurring KIM proteins, or biologically active unique fragments thereof, whose sequences differ from the wild type sequence by one or more conservative amino acid substitutions, which typically have minimal influence on the secondary structure and hydrophobic nature of the protein or

peptide. Conservative substitutions typically include the substitution of one amino acid for another with similar characteristics, such as substitutions within the following groups: valine, glycine; glycine, alanine; valine, isoleucine; aspartic acid, glutamic acid; asparagine, glutamine; serine, threonine; lysine, arginine; and phenylalanine, tyrosine.

15 Other conservative substitutions are those which meet the criteria for an "accepted point mutation" as defined in the Atlas of Protein Sequence and Structure (Dayhoff et al., eds.), 1978. See also PCT publication no. WO97/44460.

Substitution variants can also have sequences which differ by one or more non-conservative amino acid substitutions, provided however that the substitution does not  
10 abolish the native KIM protein's biological activity, ligand- or receptor-binding characteristics, or other functional property of interest. Exemplary non-conservative substitutions are those in which: (i) a hydrophilic residue, e.g., serine or threonine, is substituted for (or by) a hydrophobic residue, e.g., leucine, isoleucine, phenylalanine, or alanine; (ii) a cysteine residue is substituted for (or by) any other residue; (iii) a residue  
15 having an electropositive side chain, e.g., lysine, arginine or histidine, is substituted for (or by) a residue having an electronegative charge, e.g., glutamic acid or aspartic acid; or (iv) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having such a side chain, e.g., glycine.

Other KIM variants of this invention include truncation variants (comprising at  
20 least a unique fragment of the corresponding KIM), insertion variants, and splice variants.

Other KIM variants within the invention are those with modifications which increase polypeptide stability. Such variants can contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the peptide sequence. Also included are: variants that include residues other than naturally occurring L-amino acids, such as D-  
25 amino acids or non-naturally occurring or synthetic amino acids such as beta or gamma amino acids and cyclic variants. Incorporation of D- instead of L-amino acids into the polypeptide may increase its resistance to proteases. See, e.g., U.S. Patent 5,219,990.

KIM variants can be naturally-occurring, or produced through synthetic or molecular engineering techniques. Those of skill in the art will understand and appreciate that an engineered KIM variant can provide advantageous properties, e.g., in facilitating purification, improving stability, modulating a biological function, or the like. Thus, in some instances, a KIM variant will be desired that lacks a glycosylation site, or that has decreased aggregation potential due to elimination of a hydrophobic surface, or that has improved folding efficiency due to elimination of a cysteine residue, or the like.

In general, KIM variants have at least fifty (50) % amino acid sequence homology or similarity with the corresponding KIM protein. Preferably, the variants have at least sixty-five (65) % sequence similarity, more preferably at least eighty (80) % similarity with the corresponding KIM protein. Still more preferably, the sequence similarity is at least ninety (90) %, or most preferably, at least ninety-five (95) %. Other preferred KIM variants are those which 1) share at least forty (40) % similarity to the corresponding KIM protein, and 2) share at least eighty (80) % of aligned cysteine residues with the corresponding KIM protein.

Just as it is possible to replace substituents of the amino acid backbone or scaffold, it is also possible to replace, modify or add non-amino acid moieties to the scaffold. Such moieties can occur naturally in a given KIM as a result of post-translational processing, including acetylation, methylation, phosphorylation, carboxylation or glycosylation. As desired, such moieties can be removed or added by conventional synthetic or biochemical techniques. Further, non-natural moieties can be added to produce a derivative of a KIM or KIM variant of this invention. For example, polyethylene glycol (PEG) can be linked to a KIM to improve its stability or pharmacokinetic properties.

As will be readily appreciated, anti-KIM and anti-KIM variant antibodies can be produced by conventional techniques. Specifically contemplated are polyclonal and monoclonal antibodies, including antigen-binding fragments thereof and engineered derivatives thereof. Antigen-binding fragments of intact antibodies include complete Fab fragments,  $F(ab')_2$  compounds,  $V_H$  regions, and  $F_V$  regions. Engineered derivatives of antibodies of the invention include single chain antibodies (see, e.g., WO 96/23071), as

well as human, humanized, primatized, or chimeric antibodies (see, e.g., PCT/US 95/16709). Engineered derivatives of the present antibodies are produced, generally, using standard recombinant DNA techniques (Winter and Milstein, Nature 349: 293-99, 1991). These include "chimeric" antibodies, in which the antigen binding domain from an animal antibody is linked to a human constant domain. In this instance, an antibody with the desired KIM-binding selectivity is derived initially from a nonhuman mammal (e.g., a mouse, rat or hamster), and subjected to recombinant DNA manipulation to replace all or part of the hinge and constant regions of the heavy chain and/or the constant region of the light chain, with corresponding regions from a human immunoglobulin light chain or heavy chain. (See, e.g., Cabilly et al., United States Pat. No. 4,816,567; Morrison et al., Proc. Natl. Acad. Sci. 81: 6851-55, 1984). Chimeric antibodies reduce the immunogenic responses elicited by animal antibodies when used in human clinical treatments.

In addition, recombinant "humanized" antibodies can be synthesized. Humanized antibodies are antibodies initially derived from a nonhuman mammal in which recombinant DNA technology has been used to substitute some or all of the amino acids not required for antigen binding with amino acids from corresponding regions of a human immunoglobulin light or heavy chain. That is, they are chimeras comprising mostly human immunoglobulin sequences into which the regions responsible for specific antigen-binding have been inserted (see, e.g., PCT patent application WO 94/04679). Animals are immunized with the desired antigen, the corresponding antibodies are isolated and the portion of the variable region sequences responsible for specific antigen binding are removed. The animal-derived antigen binding regions are then cloned into the appropriate position of the human antibody genes in which the antigen binding regions have been deleted. Humanized antibodies minimize the use of heterologous (inter-species) sequences in antibodies for use in human therapies, and are less likely to elicit unwanted immune responses. Primatized antibodies can be produced similarly.

If desired, fully human antibodies with KIM-binding specificity which can be produced in nonhuman animals, such as transgenic animals harboring one or more human

immunoglobulin transgenes. Such animals may be used as a source for splenocytes for producing hybridomas, as is described in U.S. 5,569,825.

### **Identification of Specific KIMs of the Invention**

To assess the dynamic biological processes of response to injury and tissue repair, a kidney ischemia-reperfusion model, which simulates acute renal failure, was employed. The results of studies in this model system now have been published, and appear in Ichimura et al. (1998), J. Biol. Chem. 273:4135-4142, the teachings of which are incorporated by reference herein. In this model, the kidney has the capacity for cell renewal (both structurally and functionally) after injury to tubular epithelial cells. The nephron is damaged functionally by an ischemic reperfusion injury that results in regional areas of proximal tubule cell death. During the repair process, the kidney proximal tubule epithelium undergoes a complex series of events including (1) cell death and cast formation in the tubule lumen (casts are aggregates of dead, semiviable and viable cells, as well as of cell debris); (2) proliferation of surviving proximal tubule epithelial cells; (3) formation of a poorly differentiated regenerative epithelium over the denuded basement membrane (this simplified epithelium expresses vimentin, a mesenchymal marker); and (4) differentiation of the regenerative epithelium to form fully functional proximal tubule epithelial cells. Gene expression analysis of the kidney at various timepoints following the onset of ischemia revealed the upregulation of many KIMs anticipated to be involved in the injury and repair/regeneration processes.

#### **1. Generation of ischemic and normal rat adult kidneys**

Ischemic injured rat kidneys are generated as described by Witzgall et al. (J. Clin. Invest. 93: 2175-2188, 1994). Briefly, the renal artery and vein from one kidney of an adult Sprague-Dawley rat are clamped for 40 minutes and then reperused. Injured kidneys are harvested from the rats at 24 hours and at 48 hours after reperfusion. Kidneys from sham-operated, normal adult Sprague-Dawley rats are also harvested.

#### **2. mRNA isolation**

Post-ischemic rat kidneys are prepared as described by Witzgall et al. (J. Clin. Invest. 93: 2175-2188, 1994). Briefly, the renal artery and vein of the left kidney of an adult Sprague Dawley rat are clamped for 40 minutes and then reperused. The ischemic kidneys are removed from rats either 24 hours or 48 hours after reperfusion. Normal rat kidneys are used as controls.

Total RNA is isolated with Trizol reagent (BRL catalog No. 1559-026). The kidney capsules are removed, then the kidneys are rinsed in PBS and immediately homogenized in Trizol solution. In some instances, the kidney is chopped into several pieces in PBS, frozen quickly in liquid nitrogen and kept at -70°C until it is processed for total RNA isolation.

The mRNA is purified from total RNA according to the manufacturer's instructions, using an mRNA purification kit (Pharmacia, Catalog No. 27-9258-02).

3(i). Representational difference analysis (RDA) method to clone short-fragment cDNAs up-regulated by ischemia

The RDA method of Hubank and Schatz (Nucleic Acid Research 22:5640-48, 1994) is performed as described, with modifications as follow. Double-stranded cDNA is synthesized using the Superscript Choice® system (BRL Catalog No. 18090), from mRNA isolated from rat kidney either 24 hours or 48 hours post-ischemia as a tester and normal rat kidney as a driver. The tester and driver cDNA is digested with DpnII and ligated to R-Bgl-12/24 oligonucleotides. The adapter-ligated cDNA is amplified by PCR, and the PCR product is digested with DpnII to remove the oligonucleotides. The tester cDNA is ligated to J-Bgl-12-24 oligonucleotides.

Full length cDNA clones for Kim-1 and NMB contained in a plasmid vector are provided by T. Ichimura. They are digested with Sau3A1, purified and used as quenching reagents. cDNA fragments, 2-3 containing a fragment of annexin II, and 3-8 containing a fragment of alpha 2 macroglobulin (both provided by T. Ichimura) are amplified by PCR with R-Bgl-24 primer, digested with DpnII to remove the oligonucleotides and used as quenching reagents. Tester cDNA is hybridized to an excess driver cDNA including the four quenching reagents described above. After mung bean nuclease digestion.

hybridization mixture is amplified by PCR to enrich the cDNA fragments upregulated by ischemia. Each of the PCR steps is tested with a range of template DNA concentrations, and the concentration giving the best signal is chosen for the next steps. The DNA mixture containing tester and driver DNA for the subtractive hybridization step is washed twice  
5 with 70% EtOH by spinning for 3 minutes.

The second round cDNA subtraction is performed by hybridizing the first round cDNA subtraction product ligated to N-Bgl-12-24 oligonucleotides to an excess of the same driver cDNA used for the first round subtraction. The third round cDNA subtraction is performed by hybridizing the second round cDNA subtraction product ligated to J-Bgl-  
10 12/24 oligonucleotides to an excess of the same driver cDNA used for the first round of cDNA subtraction at a ratio of 1:14,000. The final PCR product after the third round of subtraction is digested with DpnII to remove the oligonucleotides and ligated to pCR-Script SK(+) vector. This represents a selected cDNA library by RDA.

15 3(ii). Suppression subtractive hybridization (SSH) method to clone short-fragment cDNAs up-regulated by ischemia

SSH is performed according to the manufacturer's instructions (Clontech, Catalog No. K1804-1) with certain modifications. Briefly, double-stranded cDNA are synthesized from mRNA isolated from rat kidney 24 hours or 48 hours post-ischemia as a tester, and from normal rat kidney as a driver with the Superscript Choice™ system for cDNA  
20 synthesis (BRL Catalog No. 18090). The cDNA is synthesized according to the manufacturer's instruction except that the first strand cDNA is synthesized at 42°C. cDNA is digested with RSAI. Tester cDNA after RSAI digestion is ligated to adaptors. For the first hybridization, tester cDNA with adaptors is hybridized with excess driver cDNA at 68°C for 12 hours. For the second hybridization, the mixture is hybridized at 68°C for 22  
25 hours. The cDNA mixture after the second hybridization is initially amplified by PCR for 27 cycles only. Then, a portion of the PCR product is re-amplified for 10 or 13 cycles. The final subtracted PCR product is digested by RSAI to remove the adaptor, and separated on a gel. The three bands and the remaining smear are cut out separately and cloned to pCR-

Script SK (+) vector (Stratagene, Catalog No. 211188). This represents a selected cDNA library by SSH.

#### 4. Isolation and Characterization of Subtracted Amplified cDNA Clones

Colonies are randomly picked from the selected RDA and SSH cDNA libraries, and plasmid DNA is isolated with Qiagen plasmid kit. Each cDNA clone, referred to as a Subtracted Amplified Clone (SAC) is sequenced with the vector primers. DNA sequences are checked against GeneBank/EMBL databases for homology by the program BLAST™.

Southern blots are prepared with the initial PCR products generated from normal and injured rat kidneys. The blots are hybridized to the inserts isolated from the SACs to confirm which SACs are induced by ischemia.

#### 4. Isolation and Characterization of potential full length cDNA clones

4.25 ug of polyA+ mRNA isolated from rat kidney 24 hours post-ischemia, which has been purified twice by oligo dT chromatography, is used to generate a cDNA library. In order to obtain double-stranded cDNA with EcoRI adaptors, a Superscript Choice® system for cDNA synthesis (BRL Catalog No. 18090) is used following the supplier's protocol, except that the cDNA synthesis reaction is carried out at 42°C. The DNA is ethanol precipitated, washed, resuspended in 5 ul H<sub>2</sub>O, and ligated to 2 ug EcoRI digested and CIAP-treated Lambda ZAPII (Stratagene catalog No. 236612). The ligated DNA is packaged and used to infect E. coli XL-1 Blue MRF. The complexity of the library is 3.4x10<sup>6</sup> independent recombinant plaques. Another cDNA library is generated from 4 ug of polyA+ RNA from rat kidney 48 hours post-ischemia as described above.

The two cDNA libraries described above are screened with inserts from the SAC clones. pBluescript plasmid vector containing cDNA inserts are excised from Lambda ZAPII vector by in vivo excision. Inserts from the longest cDNA clones, referred to as Potential Full Length (PFL) clones, are sequenced. DNA sequences are checked by the program BLAST™ against the GenBank/EMBL databases for DNA homology. Predicted protein coding regions (open reading frames, or ORFs) from some of the PFL clones also were used to search the GenBank/EMBL databases.



### Diagnostic Uses of the Compounds of the Invention

Anti-KIM antibodies of the invention, which specifically bind to a protein of the invention or a unique fragment thereof, are useful in several diagnostic methods. These agents can be labeled with detectable markers, such as fluoroscopically or radiographically opaque substances, and administered to an individual to allow imaging of tissues which express KIM protein. The agents can also be bound to (conjugated to) substances, such as horseradish peroxidase, which can be used as immunocytochemical stains to allow visualization of areas of KIM protein-positive cells on histological sections. A specific antibody could be used alone in this manner, and sites where it is bound can be visualized in a sandwich assay using an anti-immunoglobulin antibody which is itself bound to a detectable marker.

Specific antibodies to KIM protein are also useful in immunoassays to measure KIM presence or concentration in samples of body tissues and fluids. Such concentrations may be correlated with different disease states. As an embodiment of particular interest, the invention includes a method of diagnosing renal injury, or of monitoring a process of renal repair, by measuring the concentration of KIM or of KIM fragments in the urine, plasma or serum of a patient. Similarly, KIM can be measured in urine sediment, in particular in cellular debris in the urine sediment. Casts of renal tubule cells, which may be present in urine sediment from patients with ongoing renal disease, may contain elevated levels of KIM protein and mRNA.

Specific antibodies to KIM protein may also be bound to solid supports, such as beads or dishes, and used to remove the ligand from a solution, either for measurement, or for purification and characterization of the protein or its attributes (such as post-translational modifications). Such characterization of an individual's KIM protein is expected to be useful in identifying deleterious mutants or processing defects which interfere with KIM function and are associated with abnormal phenotypes. Each of these techniques is routine to those of skill in the immunological arts.

Additional imaging methods utilize KIM or KIM fragments, fused to imageable moieties, for diagnostic imaging of tissues that express KIM ligands, particularly tumors.

Further diagnostic techniques are based on demonstration of upregulated KIM mRNA in tissues, as an indication of injury-related processes. (See, e.g., PCT publication no. WO 97/44460).

#### Therapeutic Uses of the Compounds of the Invention

5           The therapeutic methods of the invention involve selectively promoting or inhibiting cellular responses that are dependent on the presence or concentration of any KIM, including without limitation KIM ligation, KIM enzyme activity, KIM secretion, KIM signalling, and KIM modulation of gene expression. In instances where a KIM and a KIM ligand are both membrane bound, and expressed by different cells, the signal  
10 transduction may occur in the KIM-expressing cell, in the KIM ligand-expressing cell, or in both.

          KIM ligation-triggered response in a KIM ligand-expressing cell may be generated by contacting the cell with exogenous KIM, KIM fusion proteins or activating antibodies against KIM ligand, either in vitro or in vivo. Further, responses of the KIM ligand-  
15 expressing cell that would otherwise be triggered by endogenous KIM could be blocked by contacting the KIM ligand-expressing cell with a KIM ligand antagonist (e.g., an antagonist antibody that binds to KIM ligand), or by contacting the endogenous KIM with an anti-KIM antibody or other KIM-binding molecule which prevents the effective ligation of KIM with a KIM ligand.

20           Similarly, the responses triggered by KIM ligation in the KIM-expressing cell may be promoted or inhibited with exogenous compounds. For example, KIM ligation-triggered response in a KIM-expressing cell may be generated by contacting the cell with a soluble KIM ligand, or certain anti-KIM activating antibodies. Further, responses of the KIM-expressing cell that would otherwise be triggered by interaction with endogenous  
25 KIM ligand could be blocked by contacting the KIM-expressing cell with an antagonist to KIM (e.g., a blocking antibody that binds to KIM in a manner that prevents effective, signal-generating KIM ligation), or by contacting the endogenous KIM ligand with an anti-

KIM ligand antibody or other KIM ligand-binding molecule which prevents the effective ligation of KIM with the KIM ligand.

Which of the interventions described above are useful for particular therapeutic uses depend on the relevant etiologic mechanism of either the pathologic process to be inhibited, or of the medically desirable process to be promoted, as is apparent to those of skill in the medical arts. For example, where KIM ligation results in desirable cellular growth, maintenance of differentiated phenotype, resistance to apoptosis induced by various insults, or other medically advantageous responses, one of the above-described interventions that promote ligation-triggered response may be employed. In the alternative, one of the inhibitory interventions may be useful where KIM ligation invokes undesirable consequences, such as neoplastic growth, deleterious loss of cellular function, susceptibility to apoptosis, or promotion of inflammation events.

Following are examples of the previously described therapeutic methods of the invention. One therapeutic use of the KIM-related compounds of the invention is for treating a subject with renal disease, promoting growth of new tissue in a subject, or promoting survival of damaged tissue in a subject, and includes the step of administering to the subject a therapeutically effective amount of a KIM protein of the invention, or of a pharmaceutical composition which includes a protein of the invention. The protein used in these methods can be a unique fragment of a full-length KIM protein, a soluble KIM variant or a soluble ligand thereof, a KIM fusion protein, or a KIM agonist. These methods can also be practiced by administering to the subject a therapeutically effective amount of an agonist antibody of the invention, or a pharmaceutical composition which includes an agonist antibody of the invention. A KIM protein can be administered concurrently with a therapeutically effective amount of a second compound which exerts a medically desirable adjunct effect. While tissues of interest for these methods include any tissue, preferred tissues include renal tissue, liver, neural tissue, heart, stomach, small intestine, spinal cord, or lung. Particular renal conditions which are expected to be beneficially treated with the compounds of the invention include acute renal failure, acute nephritis, chronic renal failure, nephrotic syndrome, renal tubule defects, kidney

transplants, toxic injury, hypoxic injury, and trauma. Renal tubule defects include those of either hereditary or acquired nature, such as polycystic renal disease, medullary cystic disease, and medullary sponge kidney. This list is not limited, and may include many other renal disorders (see, e.g., Harrison's Principles of Internal Medicine, 13th ed., 1994, which is herein incorporated by reference.) The subject of the methods may be human.

A therapeutic intervention for inhibiting growth of undesirable, KIM- or KIM-ligand-expressing tissue in a subject includes the step of administering to the subject a therapeutically effective amount of a KIM antagonist (e.g., an antibody blocks KIM ligation). In an embodiment of interest, the KIM antagonist or anti-KIM antibody can be used therapeutically to inhibit or block growth of tumors which depend on KIM protein for growth.

Other methods of the invention include killing KIM ligand-expressing tumor cells, or inhibiting their growth, by contacting the cells with a fusion protein of a KIM and a toxin or radionuclide, or an anti-KIM ligand antibody conjugated to a toxin or radionuclide. The cell can be within a subject, and the protein or the conjugated antibody is administered to the subject.

Also encompassed within the invention is a method for targeting a toxin or radionuclide to a cell expressing a KIM, comprising contacting the cell with a fusion protein comprising a KIM ligand and a toxin or radionuclide, or an anti-KIM antibody conjugated to a toxin or radionuclide. Another embodiment includes the method of suppressing growth of a tumor cell which expresses KIM, comprising contacting the cell with a fusion protein of KIM ligand and a toxin or radionuclide or with an anti-KIM antibody conjugated to a toxin or radionuclide; the cell may be within a subject, and the protein administered to the subject.

The term "subject" as used herein is taken to mean any mammal to which KIM can be administered. Subjects specifically intended for treatment with the method of the invention include humans, as well as nonhuman primates, sheep, horses, cattle, goats, pigs, dogs, cats, rabbits, guinea pigs, hamsters, gerbils, rats and mice.

### Use of Compounds of the Invention in Gene Therapy

The KIM genes of the invention are introduced into damaged tissue, or into tissue where stimulated growth is desirable. Such gene therapy stimulates production of KIM protein by the transfected or transformed cells, promoting cell growth and/or survival of cells that express the KIM protein.

In a specific embodiment of a gene therapy method, a gene coding for a KIM protein is be introduced into a renal target tissue. The KIM protein is expected to be stably expressed and stimulate tissue growth, division, or differentiation, or to potentiate cell survival. Furthermore, a KIM gene can be introduced into a target cell using a variety of well-known methods that use either viral or non-viral based strategies.

Once introduced into a target cell, sequences of interest can be identified by conventional methods such as nucleic acid hybridization using probes comprising sequences that are homologous/complementary to the inserted gene sequences of the vector. In another approach, the sequence(s) may be identified by the presence or absence of a "marker" gene function (e.g, thymidine kinase activity, antibiotic resistance, and the like) caused by introduction of the expression vector into the target cell.

### Formulation

In general, compounds of the invention are suspended, dissolved or dispersed in a pharmaceutically acceptable carrier or excipient. The resulting therapeutic composition does not adversely affect the subject's homeostasis, particularly electrolyte balance. Thus, an exemplary carrier comprises normal physiologic saline (0.15M NaCl, pH 7.0 to 7.4). Other acceptable carriers are well known in the art and are described, for example, in Remington's Pharmaceutical Sciences, Gennaro, ed., Mack Publishing Co., 1990. Acceptable carriers can include biocompatible, inert or bioabsorbable salts, buffering agents, oligo- or polysaccharides, polymers, viscosity-improving agents, preservatives, and the like. In some embodiments, the term "carrier" encompasses liposomes and the HIV-1 tat protein (See Chen et al., Anal. Biochem. 227: 168-175, 1995) as well as any plasmid and viral expression vectors.

Any KIM compound of this invention can be used in the form of a pharmaceutically acceptable salt. Suitable acids and bases which are capable of forming salts with the polypeptides, nucleic acids and vectors of the present invention are well known to those of skill in the art, and include inorganic and organic acids and bases.

5 KIM compounds of the invention are dispersed in the carrier to concentrations sufficient to deliver to the subject a therapeutically effective amount of the compound, which is an amount sufficient to produce a detectable, preferably medically beneficial effect in the subject. Medically beneficial effects would include preventing, delaying or attenuating deterioration of, or detectably improving, the subject's medical condition. It is  
10 expected that the concentration or amount of a KIM compound that will produce a medically beneficial effect will vary considerably with the circumstances in which the invention is practiced. An effective amount can be determined by an ordinarily skilled physician or other practitioner through no more than routine experimentation. As an example, an indication of the status of renal injury or renal function can be monitored with  
15 one or more routine laboratory tests which measure the concentrations of relevant substances in blood or urine, other urine characteristics, or the rate of clearance of various substances from the blood into the urine. The parameters measured by these tests, either individually or in combination, can be used by a physician to assess renal function or damage. Examples of such parameters include the blood concentration of urea, creatinine  
20 or protein; the urine concentration of protein or of various blood cells such as erythrocytes or leucocytes; urine specific gravity; amount of urine; the clearance rates of inulin, creatinine, urea or p-aminohippuric acid; and the presence of hypertension or edema.

In some embodiments, a KIM compound is formulated in a liposome delivery system, including without limitation any of a variety of unilamellar vesicles, multilamellar  
25 vesicles, or stable plurilamellar vesicles, all of which can be prepared and administered according to methods well known to those of skill in the art, for example in accordance with the teachings of United States Patent 5,169,637, 4,762,915, 5,000,958 or 5,185,154. In addition, it may be desirable to express the novel polypeptides of this invention, as well as other selected polypeptides, as lipoproteins, in order to enhance their binding to

liposomes. As an example, treatment of human acute renal failure with liposome-encapsulated KIM protein may be performed in vivo by introducing a KIM protein into cells in need of such treatment using liposomes. The liposomes can be delivered via catheter to the renal artery. The recombinant KIM protein is purified, for example, from CHO cells by immunoaffinity chromatography or any other convenient method, then mixed with liposomes and incorporated into them at high efficiency. The encapsulated protein may be tested in vitro for any effect on stimulating cell growth.

#### Routes of Administration

The compounds of the invention may be administered in any manner which is medically acceptable. Depending on the specific circumstances, local or systemic administration may be desirable. Preferably, the compound is administered via a parenteral route such as by an intravascular, intravenous, intraarterial, subcutaneous, intramuscular, intratumor, intraorbital, intraventricular, intraperitoneal, subcapsular, intracranial, intraspinal, or intranasal injection, infusion or inhalation. The compound also may be administered by implantation of an infusion pump, or a biocompatible or bioerodable sustained release implant, or by installation of a catheter (e.g., in a renal artery), into the subject. Alternatively, certain compounds of the invention, or formulations thereof, may be appropriate for oral or enteral administration. Still other compounds of the invention will be suitable for topical administration.

#### Treatment Regimes

Determining appropriate dosage and frequency of treatment with any particular KIM compound to be administered to an individual is within the skills and clinical judgement of ordinary practitioners. The general dosage and treatment schedule is established by preclinical and clinical trials, which involve extensive but routine studies to determine the optimal administration parameters of the compound. Even after such recommendations are made, the practitioner will often vary these dosages for different individuals based on a variety of considerations, such as the individual's age, medical status, weight, sex, and concurrent treatment with other pharmaceuticals. Determining the

optimal dosage and administration regime for each KIM compound is a routine matter for those of skill in the pharmaceutical and medical arts.

### Equivalents

5 The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The foregoing embodiments are therefore to be considered in all respects illustrative of, rather than limiting on, the invention disclosed herein. Scope of the invention thus is indicated by the appended claims rather than by the foregoing description, and all changes which come within the meaning and range of equivalency of the claims are intended to be embraced therein.



## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Sanicola-Nadel, Michele  
Hession, Catherine A  
Wei, Henry  
Cate, Richard L
- (ii) TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
- (iii) NUMBER OF SEQUENCES: 112
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Patent Administrator, Biogen, Inc.
  - (B) STREET: 14 Cambridge Center
  - (C) CITY: Cambridge
  - (D) STATE: MA
  - (E) COUNTRY: USA
  - (F) ZIP: 02142
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 60/047,490
  - (B) FILING DATE: 23-MAY-1997
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 60/047,491
  - (B) FILING DATE: 23-MAY-1997
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Fenton, Gillian M
  - (B) REGISTRATION NUMBER: 36,508
  - (C) REFERENCE/DOCKET NUMBER: A028PCT
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: (617) 679-2810
  - (B) TELEFAX: (617) 679-2838

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2058 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW010

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 8..646

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (1618..1885)  
(D) OTHER INFORMATION: /label= SAC\_23807

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GTCGACA	GCT	CGG	ATC	CCC	CCA	TCA	CCT	ACC	TCA	CAG	ACT	GCT	CAG	ACA	49	
	Ala	Arg	Ile	Pro	Pro	Ser	Pro	Thr	Ser	Gln	Thr	Ala	Gln	Thr		
	1				5					10						
GAA	GGT	AGT	AGG	ATC	ACA	TCC	AAA	ATG	CCT	GTG	GCG	TCT	ACC	CAG	GGG	97
Glu	Gly	Ser	Arg	Ile	Thr	Ser	Lys	Met	Pro	Val	Ala	Ser	Thr	Gln	Gly	
15					20				25						30	
TCC	TCT	CCC	CCT	ACC	CCG	GCT	CTG	TGG	GCA	ACC	TCC	CCC	AGG	GCT	GCC	145
Ser	Ser	Pro	Pro	Thr	Pro	Ala	Leu	Trp	Ala	Thr	Ser	Pro	Arg	Ala	Ala	
				35				40						45		
ACC	CTA	CCT	CCC	AAG	TCC	TCC	TCC	ACC	GGC	ACA	TCC	TGT	GAA	CCA	AAA	193
Thr	Leu	Pro	Pro	Lys	Ser	Ser	Ser	Thr	Gly	Thr	Ser	Cys	Glu	Pro	Lys	
			50					55					60			
ATG	GTC	ATC	AAC	ACG	GTC	CCA	CAG	CTC	CAC	TCG	GAG	AAG	ACA	GTG	TAT	241
Met	Val	Ile	Asn	Thr	Val	Pro	Gln	Leu	His	Ser	Glu	Lys	Thr	Val	Tyr	
	65					70						75				
CTC	AAG	TCC	AGT	GAC	AAC	CGC	CTG	CTC	ATG	TCT	CTC	CTC	CTC	TTC	CTC	289
Leu	Lys	Ser	Ser	Asp	Asn	Arg	Leu	Leu	Met	Ser	Leu	Leu	Leu	Phe	Leu	
	80				85						90					
TTT	GTC	CTC	TTC	CTC	TGC	CTC	TTT	TCC	TAC	AAC	TGC	TAC	AAG	GGC	TAC	337
Phe	Val	Leu	Phe	Leu	Cys	Leu	Phe	Ser	Tyr	Asn	Cys	Tyr	Lys	Gly	Tyr	
95					100					105					110	
CTG	CCC	GGA	CAG	TGC	TTA	AAG	TTC	CGC	TCA	GCC	CTG	CTG	CTC	GCA	AAG	385
Leu	Pro	Gly	Gln	Cys	Leu	Lys	Phe	Arg	Ser	Ala	Leu	Leu	Leu	Ala	Lys	
				115					120					125		
AAA	AAA	CCT	AAG	TCA	GAG	TTC	TCT	GAC	CTG	GAG	CAG	AGT	GTG	AAG	GAG	433
Lys	Lys	Pro	Lys	Ser	Glu	Phe	Ser	Asp	Leu	Glu	Gln	Ser	Val	Lys	Glu	
		130						135					140			
ACG	CTG	GTA	GAA	CCT	GGG	AGC	TTC	TCG	CAG	CAG	AAC	GGC	GAC	CAG	CCC	481
Thr	Leu	Val	Glu	Pro	Gly	Ser	Phe	Ser	Gln	Gln	Asn	Gly	Asp	Gln	Pro	
		145					150					155				
AAG	CCA	GCC	TTG	GAT	ACC	GGC	TAT	GAA	ACC	GAG	CAG	GAC	ACT	ATC	ACC	529
Lys	Pro	Ala	Leu	Asp	Thr	Gly	Tyr	Glu	Thr	Glu	Gln	Asp	Thr	Ile	Thr	
	160					165					170					
AGC	AAG	GTC	CCC	ACC	GAT	CGA	GAG	GAC	TCG	CAA	CGT	ATC	GAC	GAG	CTC	577
Ser	Lys	Val	Pro	Thr	Asp	Arg	Glu	Asp	Ser	Gln	Arg	Ile	Asp	Glu	Leu	
175					180					185					190	

TCC GCC AGG GAC AAA CCG TTT GAT GTC AAG TGT GAA CTC AAG TTT GCA Ser Ala Arg Asp Lys Pro Phe Asp Val Lys Cys Glu Leu Lys Phe Ala 195 200 205	625
GAC TCG GAT GCC GAC GGG GAC TGAGGCCAGT GTGTCCCAGC CCGTGCCCT Asp Ser Asp Ala Asp Gly Asp 210	676
CCGTCTTCGT GGAGAGTGTG GTGTTGAACC TATTCAGTAG CCGAGTCTTG TCACTGTGCA	736
CCAGCCTCAG TCTTTTGTC CTTTCTCTCT TGGGTTGAGC CTGTGACTTG TCCCCTTTGT	796
CCTTTTGGGA AACAAGTATC TATTCAAGCC TCAAGTCCCG GCAGTTGTTG GAGCGCTTAC	856
TACACACCTG AGCCCTTTGT GTCCTGGGGG AGAGATGGCC ACCTCCGTGG GCTGAGAAGA	916
ACCACCCCTT CCTCTTCCGC TCCTCGAAGC AGCCACTGAG AGATAATTTA ATTCCAGATT	976
GGAAATGACC TTTTGGTTT ATCAGATTGG TAACTTAATC TCCTGCCATC CGGGTGGCAC	1036
GGACGTTTTT CTTTCACTTC GTTATTTTTT TTAGGATTG CGCTCCTACT GTGTTGATGT	1096
CTTAGGTCAT TTTTTTTTTT TTAAGTTACC AGAGGAGATG TTTTGATATT CATGAGAAGA	1156
AGAACATTTT CTAGATTTTT TTTTTTTTGT GTTACATATT GAGATAAAAT ATGCCTCTGT	1216
TGCTTAAGAT TCTCAGGGAT AGACGTGATT TTTGCTAATT TCTTTCCTGC TGTTCGGAAC	1276
GTAGGCCTAA AACTGTCTCT TGAGCTCACA CACTCACCTT TCCTTTTGG TGGGTTTTTT	1336
TTTTTTTTTG TTTTGTTTTA TTAATCTTAC TCATTTGAA GGATTTTCTT TATGAGCTTT	1396
TGTTTTGTTT TTGTTTTGTT TTTTGTTTT TCTCTTTTTC CTGCCGTACA TCATCTACGA	1456
GGTGCGTTTT GAGTGAGGGC AGATGGCCCA GTGGCTTCGG GTGGCCAGT GGCTGTGGGT	1516
GGCAGTTGAG CTGGTCCTGT GAGGGGAGGA GGGCTTAGAC GCCACGGCCC TGCTGCTTCT	1576
CGGCACCTCC TGTCCCGATG GGTGGCTCT TCATGAGGAT TACATCCTGT CTCTGCTGGC	1636
TTCCATCTCG CACCTCTGCC CTTGGACTTC CACCTTGACT GTCCACAAAA GACAGAAATG	1696
GGCTGGCTAG TTGGGCTCCC GGCCTCGGAC GGTGGCCGCA ACGTCCCGTG TGGGCGGCTG	1756
CCGCTGCAGC CTGACTCCTG CCGGTGTCTT TCAGGATGTG AACGGGTGGT ACGAATCTTG	1816
ACATTTGTTT CTCACCTCCT GTGTATGAAA CGACACTCAT TCCATGTAGA GGGTGACGGA	1876
CTCTGGATCC CCCCCCTTGT CGTGTAGACA CTCATCTTCA GCGTGACCTG GTCCTGCCAT	1936
TCGGTGTAAT CATTGTGTT TATAAGATTT ACTTTGTTTT TATTTTCTA CTTGGAACGT	1996
TACATATTTG AAAAAGTACC CAAATAAACC AGAAGCTTTA TCGTTGAAAA AAAAAGTCG	2056
AC	2058

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 amino acids
- (B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ala Arg Ile Pro Pro Ser Pro Thr Ser Gln Thr Ala Gln Thr Glu Gly  
 1 5 10 15  
 Ser Arg Ile Thr Ser Lys Met Pro Val Ala Ser Thr Gln Gly Ser Ser  
 20 25 30  
 Pro Pro Thr Pro Ala Leu Trp Ala Thr Ser Pro Arg Ala Ala Thr Leu  
 35 40 45  
 Pro Pro Lys Ser Ser Ser Thr Gly Thr Ser Cys Glu Pro Lys Met Val  
 50 55 60  
 Ile Asn Thr Val Pro Gln Leu His Ser Glu Lys Thr Val Tyr Leu Lys  
 65 70 75 80  
 Ser Ser Asp Asn Arg Leu Leu Met Ser Leu Leu Leu Phe Leu Phe Val  
 85 90 95  
 Leu Phe Leu Cys Leu Phe Ser Tyr Asn Cys Tyr Lys Gly Tyr Leu Pro  
 100 105 110  
 Gly Gln Cys Leu Lys Phe Arg Ser Ala Leu Leu Leu Ala Lys Lys Lys  
 115 120 125  
 Pro Lys Ser Glu Phe Ser Asp Leu Glu Gln Ser Val Lys Glu Thr Leu  
 130 135 140  
 Val Glu Pro Gly Ser Phe Ser Gln Gln Asn Gly Asp Gln Pro Lys Pro  
 145 150 155 160  
 Ala Leu Asp Thr Gly Tyr Glu Thr Glu Gln Asp Thr Ile Thr Ser Lys  
 165 170 175  
 Val Pro Thr Asp Arg Glu Asp Ser Gln Arg Ile Asp Glu Leu Ser Ala  
 180 185 190  
 Arg Asp Lys Pro Phe Asp Val Lys Cys Glu Leu Lys Phe Ala Asp Ser  
 195 200 205  
 Asp Ala Asp Gly Asp  
 210

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1727 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW011

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCGACATGA TTGCTAGTGG TTCTGCTGGC TGTGAGGGCT CCCTGTCTGT GAGATCTTGT	60
TCGGGCTTTC CACATCCAGG ATCCATCCAT TCACCCCTCA CCCACTCACC CTGTGGAGCA	120
TAAAACCTTT GCAGTCAGCC AGCATGGCAG AATGTCTATG CCTTGCTTCA GTTTCAGGG	180
TAAC TAAGCA TGTTTTAATC CCACTGCCCC ACCCACTATT GTCATTCTG TCTAGTCTTC	240
TGCTACCCAG CTTGTGTGGC TCCGTTCCCTG TCACGCCTTC AGACGTCACT GTTCTGTGAC	300
AGCGGTCTCT TTCCCCTCTG CTCCACCCAT TCTGAGCTCT GGGATAGTCC TAAGCCCTGC	360
CCAGAGTAGG TGGACTGGTC GTTCCATGTC CGTGCTTGTG GGAGTTGTGG GGGTCTGTGG	420
CATCTGCTAA GGAGGTTTAT CCTCACGTCA CGGAATCCGG CAAGATGCGC CATGGTCTAC	480
ACTGCTTTCT CGAGCAGCGG CGAGAAAGGC CAGGATAACA CACGCCTCTC TGCCCTGAGA	540
TAACAACTG TCTGAGAAAG GAAGTGGCCC GTGTCTGGAA CATTGGCCCC AGTAAAGCAG	600
AGGAGACACT GCCATGTTTC CGTCATTCTG TTCACTGCC TTGCTCTGGG GCTCAGGGGT	660
TGAAGAATGA GGATACACCA TTCGACCTTC AGGCCTTCCT CAGACCCTCT TACAAAGCCT	720
GTTGTTGTGA GTCCAGCTCA TCAGCATGGC CGCCTTGTTT ACAGTTTAGC ACCAGTGTCT	780
GCAACCGTGT CTGAAACGTA GCAGATGTCC AGTACTTTTT GTAAATGAAA AAATGAACAG	840
AGTGAAGAGA AAGAGTCACC ATGAAATTTA CTAGTCCAGT GAGCTGACCG TAGGGGCAGT	900
ATAAACATCC CCCAGCTCTT AATGTTCTGT CCTCAAGCTG TCCCCTCCCA CCTCTCCGCT	960
GGCAGGAGAG TGCTCTGTAT ATGTGGACAG TCGTTGAATG GCAGGTCCAG GATGTCCTCT	1020
GCAGGTGGAC AAAAGAGAAT TCTGCTTCCC GGAAGTCTCC TTTGCTGGTG ATCTCCATGT	1080
CCATCCTCAC TGAGCACTAG AATTAAAACC ACTCAATAAT GTGATATGTG TGGTCACCTT	1140
CAGTTCAGTA GCAGTATTTT CTGGTCATTA ATATTCTGCA TTTTATTATG GTAAGATAGT	1200
ATAAGAAAAT GGTGGGCTGG TATATAGCCA GCCATGAGAC CACTCCTTAC GTCAAGGACT	1260
GGCAGAGTCT GAGGGACCTG AGTCTGGGTC CTCCTCCTGT CCCCTCCTCC TGTAGCTATG	1320
GTGGCCTTCA GTATTCTGAG ACTGAAGGCT ACTACTTCAG GACACCTACT TGATCTTAGT	1380
ACTCTCAGAA CAGCTCTCCT CAGTTCTAGC TGTGATCGCG GGCTGCGTTT GGAACGTAA	1440
CGTACTAATG GGTTTAAGAC TTTCAGAAAG TGATTTTAAA ACAAATGAC TGGAAGTTTT	1500
GTCTGATAAA GCAAGTAGAA TGGAATGTCT CAGACTGTCT CGTCACATTT TCATGTTTCA	1560
TACGAAAAGA AAAGATTGAT AACCTATCTT TAGAAATAGT AATGTTAGTT CTTTCCTATT	1620
TTGTTTAAAC CTTGAAGAAG GTTTTATTTA ACTCGTGAAC TTGTTCTTTT GTTATGACAT	1680
GATTGTTCCA GATTTGAGGG ACAAATAAAA ATTAACCTTT CTAAAAA	1727

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW012

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 39..671

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTCGACGCAT	GGGTGCAGGT	GACAAATGAG	CTGAAGAA	ATG	ATC	CCC	AGA	GGA		53						
				Met	Ile	Pro	Arg	Gly								
				1				5								
ATA	GGA	TTG	TGG	CTG	CCT	TCA	GCT	CTG	CTC	CTC	TCT	CGG	GTC	CCA	GGC	101
Ile	Gly	Leu	Trp	Leu	Pro	Ser	Ala	Leu	Leu	Leu	Ser	Arg	Val	Pro	Gly	
				10				15						20		
TGT	TTC	CCT	CTG	CGT	GGC	CCC	AGC	AGT	GTG	ACA	GGC	ACT	GTG	GGG	GAG	149
Cys	Phe	Pro	Leu	Arg	Gly	Pro	Ser	Ser	Val	Thr	Gly	Thr	Val	Gly	Glu	
				25				30					35			
TCC	CTG	AGC	GTG	ACT	TGT	CAG	TAT	GAG	GAG	AGA	TTC	AAG	ATG	AAT	AAG	197
Ser	Leu	Ser	Val	Thr	Cys	Gln	Tyr	Glu	Glu	Arg	Phe	Lys	Met	Asn	Lys	
				40				45					50			
AAA	TAC	TGG	TGC	AGA	GGG	TCA	CAT	GTT	CTA	CTT	TGC	AAA	GAT	ATT	GTC	245
Lys	Tyr	Trp	Cys	Arg	Gly	Ser	His	Val	Leu	Leu	Cys	Lys	Asp	Ile	Val	
				55				60					65			
AAG	ACC	GGA	GGC	TCA	GAA	GAA	GCT	AGA	AAT	GGC	CGA	GTG	TCC	ATC	AGG	293
Lys	Thr	Gly	Gly	Ser	Glu	Glu	Ala	Arg	Asn	Gly	Arg	Val	Ser	Ile	Arg	
				70				75			80				85	
GAT	GAT	CCA	GAC	AAC	CTG	ACC	TTC	ACA	GTG	ACC	TTG	GAG	AGC	CTC	ATC	341
Asp	Asp	Pro	Asp	Asn	Leu	Thr	Phe	Thr	Val	Thr	Leu	Glu	Ser	Leu	Ile	
				90					95					100		
CTG	GAG	GAT	GCA	GGC	ACC	TAC	ATG	TGT	GGG	GTG	GAT	ATA	CCA	TTT	ACT	389
Leu	Glu	Asp	Ala	Gly	Thr	Tyr	Met	Cys	Gly	Val	Asp	Ile	Pro	Phe	Thr	
				105				110					115			
AAT	CAC	CCC	TTG	GGG	ATT	GAT	GAG	TTC	TTC	AAG	GTT	GAA	TTG	TCT	GTG	437
Asn	His	Pro	Leu	Gly	Ile	Asp	Glu	Phe	Phe	Lys	Val	Glu	Leu	Ser	Val	
				120				125					130			
GTC	CCA	GGT	TCA	AGC	CTT	TGG	AGC	AGT	ACA	ACA	GTC	CCA	GAG	ACC	ATC	485
Val	Pro	Gly	Ser	Ser	Leu	Trp	Ser	Ser	Thr	Thr	Val	Pro	Glu	Thr	Ile	
				135				140					145			

AGA TCC TCA CTG GTT CAT ACT CAG CCC AGT GTG ACC ACA GAA GAC ACA	533
Arg Ser Ser Leu Val His Thr Gln Pro Ser Val Thr Thr Glu Asp Thr	
150 155 160 165	
ATT CCT GCT CCT AGT GCA CGG CCT CGG TCT CTT CTG GGC AGC CTC TAC	581
Ile Pro Ala Pro Ser Ala Arg Pro Arg Ser Leu Leu Gly Ser Leu Tyr	
170 175 180	
CTC TGC ATC CTG GTG TTT CTG GAG TTG CCC CTG TTC CTG TGC ATG CTC	629
Leu Cys Ile Leu Val Phe Leu Glu Leu Pro Leu Phe Leu Cys Met Leu	
185 190 195	
TGT GCC GTC CTG TGG GTG AAC AGG CCT CAG AGG TGC AGT GGG	671
Cys Ala Val Leu Trp Val Asn Arg Pro Gln Arg Cys Ser Gly	
200 205 210	
TGACATAGCA TTGGGCCCTG TCATGAGTAC CAGTGAAGTC TGTGATATG GAGGCCCTGT	731
CCCTGGATGC ATCACCTCTG GTGGCCAAGG ACAACCAATA AAGCTGTGTT CTTGAGAATG	791
CTCTGAGACT TTTAGAAGAT TCTCTGCACG TGTCAAACAC ATAAAAGAAT GTGCCCAATA	851
AATACCAGCT GGGGAATTGG CTCAGCAGGT AACAACAATG GCTGCTTTTC CAGAGGTCCT	911
GAGTTCAATT CCAATAACTA CATCGGTGGC TCACAACCAT CTCTAATGAG TTCAGATGCC	971
CTCTTCTGCT GTGCAGGTGC ACATTCACAT AGAATATTCA TACATTGAGT AAATTTATTA	1031
AAAGTAATTG CAGAATTAAA ACACAAAAA	1060

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Ile	Pro	Arg	Gly	Ile	Gly	Leu	Trp	Leu	Pro	Ser	Ala	Leu	Leu	Leu
1				5					10					15	
Ser	Arg	Val	Pro	Gly	Cys	Phe	Pro	Leu	Arg	Gly	Pro	Ser	Ser	Val	Thr
		20						25					30		
Gly	Thr	Val	Gly	Glu	Ser	Leu	Ser	Val	Thr	Cys	Gln	Tyr	Glu	Glu	Arg
		35						40					45		
Phe	Lys	Met	Asn	Lys	Lys	Tyr	Trp	Cys	Arg	Gly	Ser	His	Val	Leu	Leu
	50					55					60				
Cys	Lys	Asp	Ile	Val	Lys	Thr	Gly	Gly	Ser	Glu	Glu	Ala	Arg	Asn	Gly
	65				70					75					80
Arg	Val	Ser	Ile	Arg	Asp	Asp	Pro	Asp	Asn	Leu	Thr	Phe	Thr	Val	Thr
				85					90					95	
Leu	Glu	Ser	Leu	Ile	Leu	Glu	Asp	Ala	Gly	Thr	Tyr	Met	Cys	Gly	Val
			100					105					110		

Asp Ile Pro Phe Thr Asn His Pro Leu Gly Ile Asp Glu Phe Phe Lys  
 115 120 125

Val Glu Leu Ser Val Val Pro Gly Ser Ser Leu Trp Ser Ser Thr Thr  
 130 135 140

Val Pro Glu Thr Ile Arg Ser Ser Leu Val His Thr Gln Pro Ser Val  
 145 150 155 160

Thr Thr Glu Asp Thr Ile Pro Ala Pro Ser Ala Arg Pro Arg Ser Leu  
 165 170 175

Leu Gly Ser Leu Tyr Leu Cys Ile Leu Val Phe Leu Glu Leu Pro Leu  
 180 185 190

Phe Leu Cys Met Leu Cys Ala Val Leu Trp Val Asn Arg Pro Gln Arg  
 195 200 205

Cys Ser Gly  
 210

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3385 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

(B) CLONE: HW013

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTGCACTTTT TTTAACAAAG ATGATTTAAT ATTTAATATC ATACCCAGAA AAATCAAACA	60
ACTTCCAGAA TCAGACAGGT AAGATTCCTT AGTATGTTTT TACCAAAATT TTATAAGATT	120
ATCTTAAATT TGAAGAATTG TGCATAAATA ACACAGGATA CCAAAGGAAA CTGAGCTGTA	180
TTTTTTGCTT TGATGGGAAA AAAAAAAAAA AAAACCTGAA ACCCAGAGAC ACTAACCATT	240
TTGTGCACTG TATAACTCAT TGTGGAAGCT GAAACTTAGA ACTCATCTTT GGAAATATAT	300
AATTGATGAA CCCACCTTTC CATTCAGCTT TTAGGGTGAT GTAAACTGT ATGAATCCAC	360
TTACAAATAA ACTTAAAAGA CCTTTTCATG TTACCCTATC TCTTATCTTC TCAATGTATT	420
TCTATATGCT TACCTTTAGT AAGAGTACCA TGGAGTACTT TGATAGATGG AATGACAGAG	480
AAATGGATGT TTAAAGATCT TTAGGATGAA AAGGATAATT TAAGAAGCAT ATTTAAGAAG	540
CATGCTGATA TTGAAAAGGT AGTATAGGCA AGTACAGTGC ACTTCCACAG ACTACTCATA	600
CCTTCTTGTC CATCAGAAAG TCTGAGTGGG TTAGTTACAG TAAGGTCATG TTGATCTGGG	660



TCTATGCACC	TGTGCTGAGG	TCAGAGGACA	ACTTTCAGGA	GTCTGTTCTG	CCCTTCCGTG	720
TGGGTCTTAA	GTATGCAACT	CAAAGGTTGT	CTGTCAGGTT	GATTCTGCCC	AGTCATGAAA	780
CCTTTTGAGAA	GAGGAAAAGG	TGTAGGAAGA	ACAGGAACAG	GATCATAGAA	GTGACTAAGC	840
CAGAATCTGT	CATATTTGAG	GAAACATCAA	GAGGTGCCTA	TAATAGGGTT	CCGAGAATGA	900
AAAAAAAAAC	CAGATAACTT	TACTGTCTAT	TCCAGCCAGT	TTGACTTAGA	GGCACTGGCA	960
AAGTAGGTAC	AGACAGCTGA	GGTAGAAAAT	GTGGAACATT	ACTGTGGCAA	CTGCTGCATC	1020
TTCGGGTGTG	GAAGTCACAC	CAATTCCACT	TTTGGTTGGC	ACTTCCCATC	TTCTAAGAGA	1080
GAACCTTTAA	TTGAGCGGCC	GTGAAGAACT	TGATAAATCT	ACGCTGCTGT	GTGTCTCTCA	1140
CCACCACCCT	ACCTACATAG	ACACCCCCAT	TTAGTTGTTT	CCTGAGAATC	AGAACAGAGC	1200
ACTTACACGA	GGTACGGGAA	GCTGTGATGC	CTAGTGTAAG	AAGTTAGAAG	ATGACTGAAC	1260
ACAGGAAGAG	AGCGGGCTTC	CAGTGGGATC	TAAATCAAAT	CAGCGAGGAT	CACTGGAAAC	1320
TGAGCTGAAA	CTCAACTAAG	ATTTAGCTTG	CAAACAATAG	AGATGTTTCA	TTTTAAAACC	1380
ACAGCAGATT	AACAGTATAG	CTGGGAAACA	ACCAAGTGTT	CTTTTCAAAT	TTCAGTTCCA	1440
CTATATCTGA	GCATGGTTAA	AATAGGGGTG	AGTCCTATTT	CATGGCTTTT	ATCTATTAAA	1500
TATTTGTAAA	CGTGCTCAAG	TTCCCATTTT	TTTAACCCTA	CAAAATTCAG	GCTTCTCACA	1560
CTGGGGCTGT	AGTTCCTTAA	TAGTTTCTA	ATTCTGAGC	CATCTTTAGA	CTTCTGAAAA	1620
CCTCAACTTC	AGTCTCTTCT	CACATAGCCA	ACTATGAGAA	CCTCCAAAGC	AGAATGCAGT	1680
TTTCCTTGTTG	CCTTTAGCAA	CAAGGCACCT	TTACAAAAC	GGTTCTCAGT	GAATGAAATG	1740
AAAGCATGAT	CTTACAGTGT	GCGACACTTC	AGTAGTGATG	ACTTAGAATA	GTTTTAAAT	1800
AAATAAAATT	ACAGAACCTA	GGGGAGGTGA	AAAATGGGAA	ACATGAAACA	GTAGACCATA	1860
CTGCAGCTCA	CAAAACAAGC	AGTGAAAACA	TATAAACTAC	CTATAATAGC	CTCATTATAA	1920
CATGCAAAACA	CACTAACTCT	GTTACATACA	GATTGGATGT	ATATCTGAAG	TTGTTTATAG	1980
ATACAATGGG	TTTAAGTTGG	TAATTAATAA	ATAGTACTAC	CTTAGAATGT	GTTGTGCTTT	2040
TAAGTATTAT	GAAAGCTCAA	GAGCCAGGCA	TCTTCCTCTT	CAGAGCTGTT	TCCTGTTACA	2100
TAATTACTGG	CAAAGTTCTA	AACAGTTCTT	TACTAAATCA	CAAAATGCAT	ACAGCTAGTG	2160
CCATTGTCTC	AAGTCTTTTT	AGAATAGAAA	TGTCGATGTG	AAGCTATAAT	TTTATTTTGG	2220
TTTTATGATC	CTGCAAGGGT	TAAGTTTGTC	TGATTCTTCC	CAGGGTACTG	ATAAGTTCTC	2280
TAACATACCG	GAGTAAGAGA	ACTGAAACAG	GTTATTTCCC	TCATCCTCAG	GAATCTGCTT	2340
GAATATGGAT	CAGCATATAT	TGGACTTAAT	GCTGCTTTTG	GGGGCCTAAT	AGCAAACAGT	2400
CTATTTCCGC	GAATTTTGAA	CGTGACACAA	GCTCGAGTAG	CATCTAGCTT	GCCAATGGCT	2460
GTGATCCCAT	TTTTGACAGC	AAACCTGTCC	TACCATAGTT	TTGTAAGTTT	ACCTTTGAGT	2520
ACAGGTATGT	TTTATTTTCAT	TATCATCCAA	GCCTGCATTT	ATGCATACTA	TATACACTTA	2580

TATATACTTC ACTCAAACT GAAGGATTAG AACCGAGTAG GTGTGAGCTT TAATGAAGTG 2640  
 TTTAGGGAAG ACCAAATCTA TTCAACAACCT GATTGAGCTT AGGTCTGCCA ACCAACCAGT 2700  
 GAGTACATAT CCACAGCCGT CAGTCACACA CTTCCCTGAC CTGAAGCAGC ATTCCCAGTG 2760  
 CTGCCTTTGT TTCTTGTTAA TTTGCACATT ATATGCAAGT TTTCTCTTTT TAAAAAGATG 2820  
 TTTTATTATA GGTATGTGTT TTGTCTGACT GTACATTTGC ACACCAGGTA GTGGTACTGG 2880  
 ACCCATATA CTAGTTACAG ATAAGGGTGA GCTGCCACGT GGGTGCTGGT AATTGAACTC 2940  
 AGCACTTAGT GCCATTGCTC TTAATTGAGA TCTCTCTCCA GCCCTGTATG AAAGTTCTTT 3000  
 AAAGTATAAG TATGTTGTGG GAATACAACCT CTGTAATCTC TGAAGTACTA TGTGCATTGC 3060  
 AGCATTTAGC AAAATGAAAG ACATAGTGTC CCATGTGTCC TGAACATGTA TTGTTATGTG 3120  
 TTCTATATGC ATGACCCCTC TTATTAAAAA CAGATTGGGG AAAGGGGTCA ATTTTGTGTT 3180  
 TTCTCCTAAA TAATAATCCA AATTTCTGAG AATCCAGAGT CAAGATTTGT TTTCTGTTT 3240  
 ATCAGTATTT CCTATCCAAG GCCTTAATAC ATGCTAAGCA TGTATTCACA CCAAGCTACA 3300  
 GTCCATCCTC TATAATTATT CCTTTGACAT GATATGCACC ATTTTGTGCTG TGCTACTGAA 3360  
 AATAAAATGT GAACCTTATG TCGAC 3385

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1779 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW014

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 175..804

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTCGACCTGG AAAGAGCTGC AGTGTTTCGCG CTTGGTAGCT GGTGGACCGG GCTCAGCACC 60  
 TTCGGCTGGC TTTGTGTCCC AGAGGCTCAC CGGAAAAAGA CTTTCTCAGC CCTCGGACTC 120  
 CAGAGAGAGA TACTTTTCGT GCTCCCCAAC CTTGACTCCT GCAGAAGCTG AGCG ATG 177  
 Met  
 1  
 GCG TCC ATG GGG CTG CAG GTG CTG GGA ATC TCC TTG GCC GTC CTG GGC 225  
 Ala Ser Met Gly Leu Gln Val Leu Gly Ile Ser Leu Ala Val Leu Gly  
 5 10 15

TGG	TTA	GGA	GTC	ATC	CTG	AGT	TGT	TCG	CTC	CCC	ATG	TGG	CGG	GTG	ACC	273
Trp	Leu	Gly	Val	Ile	Leu	Ser	Cys	Ser	Leu	Pro	Met	Trp	Arg	Val	Thr	
		20					25					30				
GCC	TTC	ATC	GGC	AGC	AAC	ATC	GTC	ACG	GCG	CAG	ACG	AGC	TGG	GAA	GGC	321
Ala	Phe	Ile	Gly	Ser	Asn	Ile	Val	Thr	Ala	Gln	Thr	Ser	Trp	Glu	Gly	
	35					40				45						
CTG	TGG	ATG	AAC	TGC	GTG	GTG	CAG	AGC	ACC	GGC	CAG	ATG	CAG	TGC	AAG	369
Leu	Trp	Met	Asn	Cys	Val	Val	Gln	Ser	Thr	Gly	Gln	Met	Gln	Cys	Lys	
	50				55					60					65	
ATG	TAC	GAC	TCG	ATG	CTG	GCC	CTG	CCG	CAG	GAC	CTG	CAG	GCC	GCC	CGA	417
Met	Tyr	Asp	Ser	Met	Leu	Ala	Leu	Pro	Gln	Asp	Leu	Gln	Ala	Ala	Arg	
				70				75						80		
GCC	CTG	ATG	GTC	ATC	AGC	ATC	ATC	GTG	GGT	GCT	CTG	GGG	ATG	CTT	CTC	465
Ala	Leu	Met	Val	Ile	Ser	Ile	Ile	Val	Gly	Ala	Leu	Gly	Met	Leu	Leu	
			85					90					95			
TCA	GTC	GTA	GGG	GGC	AAG	TGC	ACC	AAC	TGC	ATG	GAG	GAC	GAG	ACC	GTC	513
Ser	Val	Val	Gly	Gly	Lys	Cys	Thr	Asn	Cys	Met	Glu	Asp	Glu	Thr	Val	
		100					105					110				
AAG	GCC	AAG	GTC	ATG	ATC	ACA	GCC	GGA	GCC	GTG	TTC	ATC	GTG	GCA	AGC	561
Lys	Ala	Lys	Val	Met	Ile	Thr	Ala	Gly	Ala	Val	Phe	Ile	Val	Ala	Ser	
	115					120				125						
ATG	CTG	ATT	ATG	GTG	CCT	GTG	TCC	TGG	ACG	GCA	CAC	AAC	GTC	ATC	CGC	609
Met	Leu	Ile	Met	Val	Pro	Val	Ser	Trp	Thr	Ala	His	Asn	Val	Ile	Arg	
	130				135					140					145	
GAC	TTC	TAC	AAC	CCT	CTG	GTG	GCT	TCC	GGG	CAG	AAA	AGG	GAG	ATG	GGG	657
Asp	Phe	Tyr	Asn	Pro	Leu	Val	Ala	Ser	Gly	Gln	Lys	Arg	Glu	Met	Gly	
			150						155					160		
GCC	TCG	CTT	TAC	ATC	GGC	TGG	GCG	GCT	TCT	GGG	CTG	CTG	CTC	CTG	GGA	705
Ala	Ser	Leu	Tyr	Ile	Gly	Trp	Ala	Ala	Ser	Gly	Leu	Leu	Leu	Leu	Gly	
			165					170					175			
GGA	GGC	CTC	CTC	TGC	TGC	AAT	TGC	CCA	CCT	CGC	CGC	AAC	GAA	AAG	CCC	753
Gly	Gly	Leu	Leu	Cys	Cys	Asn	Cys	Pro	Pro	Arg	Arg	Asn	Glu	Lys	Pro	
		180					185					190				
TAC	TCC	GCC	AAG	TAC	TCC	GCC	GCC	CGC	TCT	GTT	CCC	GCC	AGC	AAC	TAT	801
Tyr	Ser	Ala	Lys	Tyr	Ser	Ala	Ala	Arg	Ser	Val	Pro	Ala	Ser	Asn	Tyr	
	195					200					205					
GTG	TAAGGTGGGC	CACTCTCTGT	CCACATTGCC	TTTATTTTCT	TGGATTGAAC											854
Val																
	210															
TCATAACGGC	CTGTGGCCCC	TCACATTCTC	CAGGACCTGA	CCAGCTGTGG	GCTACTGACT											914
GCTTGCAAAC	CCGGACTGTG	CTAAGTTACT	AGCGTGTAGC	CCTTGGGGAC	CCACCTGGCC											974
CATCTGGACA	CATCTCAAGG	CTCCAGCGAG	GATAGATGTA	AAAATATTTT	CTTGCTTGCA											1034
TCCAGATTGC	TCATGGATAC	GGGGCTGAAG	GCAGAAGCAG	CTGTCTGGGT	ACGACAGTGG											1094
AGGGGGAGCT	GGGTCTTGCT	GGCCGGGATA	GCTCAGCTGT	GACTTTGGTC	TCTGGAGTGG											1154

```

ATGTCCTGGT CATGTTAGCA AACATTCACT GCCCTTTCTC AGTGCCCTCG CTCTCTCGCC 1214
TCCACGTTAC TCCCGCGCTA CTCTTGCCGT TTCTCGCCCCG CGTTTCTGAG CACACCAGGT 1274
CCTGCCTGGA GTCTTGGTGT CGAGGATGAC TGA CTGAAGG GGCCTTTGAG AGCTGATGGG 1334
TTCTGCCATG GACTCCTCCC GGTGATTAGC AATGACTGGG GCCTTACCCA CCCACCTACC 1394
CTCGTAATGA AGTTCTGTGG AGTGGCTGGA CAGGTTTGAG GGAAGGGTGG AGGTGGTTTA 1454
AACTGGTTTG GGGAGTGCTA GGGCTGGGGA CCCAGAAGCA GCCCAGGGTG TCCCCACCCC 1514
TTTCCCATAC GGTCTTGCTA AATGTTCTGA TCTCTGTATA CCCCCTCCCT CTTCAGAAGG 1574
ACCCTGGGTG GGCCCTCTG AATTCCTAC CCTTGTCAT TTCAAGGACG CTGGCCAGTC 1634
TGTGGAAGGT ACGGGGTCT GATGGCATTG CACCAGGGAG CCTCCTGGAC TCCCTTGCCT 1694
TCTCTGTGGT TTCTTGTTTT GTAATTTAAG GTCTGTTCAC AGCTGTAATT ATTATTTTCT 1754
ACAATAAATG GCACCTGCAG TCGAC 1779

```

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Ala Ser Met Gly Leu Gln Val Leu Gly Ile Ser Leu Ala Val Leu
 1             5             10             15
Gly Trp Leu Gly Val Ile Leu Ser Cys Ser Leu Pro Met Trp Arg Val
 20             25             30
Thr Ala Phe Ile Gly Ser Asn Ile Val Thr Ala Gln Thr Ser Trp Glu
 35             40             45
Gly Leu Trp Met Asn Cys Val Val Gln Ser Thr Gly Gln Met Gln Cys
 50             55             60
Lys Met Tyr Asp Ser Met Leu Ala Leu Pro Gln Asp Leu Gln Ala Ala
 65             70             75             80
Arg Ala Leu Met Val Ile Ser Ile Ile Val Gly Ala Leu Gly Met Leu
 85             90             95
Leu Ser Val Val Gly Gly Lys Cys Thr Asn Cys Met Glu Asp Glu Thr
 100            105            110
Val Lys Ala Lys Val Met Ile Thr Ala Gly Ala Val Phe Ile Val Ala
 115            120            125
Ser Met Leu Ile Met Val Pro Val Ser Trp Thr Ala His Asn Val Ile
 130            135            140
Arg Asp Phe Tyr Asn Pro Leu Val Ala Ser Gly Gln Lys Arg Glu Met
 145            150            155            160

```

(2) INFORMATION FOR SEO ID NO:9:

(A) LENGTH: 2537 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(B) CLONE: HW015

(A) NAME/KEY: CDS  
(B) LOCATION: 11..2176  
(D) OTHER INFORMATION: /product= "predicted polypeptide"

GTCGACCAAA	ATG	GCA	GCC	TTG	GAA	GAA	GAA	TTC	ACG	TTG	TCT	ACT	GGA			49
	Met	Ala	Ala	Leu	Glu	Glu	Glu	Phe	Thr	Leu	Ser	Thr	Gly			
	1				5					10						
GTC	TTA	GGC	GCC	GGG	CCT	GAA	GGA	TTC	CTA	GGT	GTA	GAG	CCC	AGT	GAC	97
Val	Leu	Gly	Ala	Gly	Pro	Glu	Gly	Phe	Leu	Gly	Val	Glu	Pro	Ser	Asp	
	15					20					25					
AAG	GCC	GAC	CAG	TTC	CTA	GTG	ACC	GAC	AGA	GGC	AGG	ACC	GTC	GTC	CTC	145
Lys	Ala	Asp	Gln	Phe	Leu	Val	Thr	Asp	Arg	Gly	Arg	Thr	Val	Val	Leu	
	30				35					40					45	
TAC	AAG	GTT	TCT	GAC	CAG	AAA	CCC	TTG	GGC	AGC	TGG	TCG	GTG	AAA	CAA	193
Tyr	Lys	Val	Ser	Asp	Gln	Lys	Pro	Leu	Gly	Ser	Trp	Ser	Val	Lys	Gln	
				50					55					60		
GGC	CAG	ACT	ATA	ACA	TGT	CCT	GCT	GTG	TGC	AAC	TTT	CAG	ACT	GGA	GAG	241
Gly	Gln	Thr	Ile	Thr	Cys	Pro	Ala	Val	Cys	Asn	Phe	Gln	Thr	Gly	Glu	
			65					70					75			
TAT	ATT	ATG	GTA	CAT	GAC	CAT	AAG	GTT	TTG	AGA	ATA	TGG	AAT	AAT	GAA	289
Tyr	Ile	Met	Val	His	Asp	His	Lys	Val	Leu	Arg	Ile	Trp	Asn	Asn	Glu	
		80					85					90				
GAC	GTA	AAC	CTG	GAT	AAA	GTA	TTC	AAA	GCT	ACA	CTG	TCA	GCT	GAG	GTC	337
Asp	Val	Asn	Leu	Asp	Lys	Val	Phe	Lys	Ala	Thr	Leu	Ser	Ala	Glu	Val	

95	100	105	
CAT AGG ATC CAC TCA GTA CAA AGA ACA GAA CCC CTG GTG CTG TTC CGA His Arg Ile His Ser Val Gln Arg Thr Glu Pro Leu Val Leu Phe Arg 110 115 120 125			385
AGA GGC GCT GCT CGT GGG TTA GAG GCC TTG CTT GCA GAG CCA CAG CAG Arg Gly Ala Ala Arg Gly Leu Glu Ala Leu Leu Ala Glu Pro Gln Gln 130 135 140			433
AAC ATT GAA GCT GTC ATA CCA GAT GAG GAA GTG ATC GAG TGG TCA GAG Asn Ile Glu Ala Val Ile Pro Asp Glu Glu Val Ile Glu Trp Ser Glu 145 150 155			481
GTT TTC ATG TTA TTT AAG CAA CCG GTT TTA ATT TTT ATT ACT GAA AAT Val Phe Met Leu Phe Lys Gln Pro Val Leu Ile Phe Ile Thr Glu Asn 160 165 170			529
CAT GGG AGT TAT GTT GCT TAT GTA CAA TTG TGC AAA TCA CAC AGC TTA His Gly Ser Tyr Val Ala Tyr Val Gln Leu Cys Lys Ser His Ser Leu 175 180 185			577
AGC AAA TAC ATA CTC TTA CTT GGA AAA GAA GAA AAA TCT GCT AAA CCA Ser Lys Tyr Ile Leu Leu Leu Gly Lys Glu Glu Lys Ser Ala Lys Pro 190 195 200 205			625
AAT TTT ACT GCA CGT GTG GAT GGG AAA TTC ATC TCC CTG GTG TCG CTA Asn Phe Thr Ala Arg Val Asp Gly Lys Phe Ile Ser Leu Val Ser Leu 210 215 220			673
GGC TCT GAT GGG TGT GTA TAT GGA ACC TTG GTA CCA ATA TAT TCA AGT Gly Ser Asp Gly Cys Val Tyr Gly Thr Leu Val Pro Ile Tyr Ser Ser 225 230 235			721
GAC ACG GAA AAC AAT CAG AGG TTA GTT AGA GCA TTG ATG CTC AAG ACG Asp Thr Glu Asn Asn Gln Arg Leu Val Arg Ala Leu Met Leu Lys Thr 240 245 250			769
GTT GTG TCT GGC AGT GCT CGA AAT GGT TCT GCC CTC ACC ATC CTG GAT Val Val Ser Gly Ser Ala Arg Asn Gly Ser Ala Leu Thr Ile Leu Asp 255 260 265			817
CAA GAC CAC ATA GCA GTC CTG GGA CCT CCA CTT CCA GCT TCT AAG GAA Gln Asp His Ile Ala Val Leu Gly Pro Pro Leu Pro Ala Ser Lys Glu 270 275 280 285			865
TGC CTC TCC ATA TGG AAC ATA AAA TTT CAG ACA TTA CAG ACA TCA AAA Cys Leu Ser Ile Trp Asn Ile Lys Phe Gln Thr Leu Gln Thr Ser Lys 290 295 300			913
GAG CTG CCA CAA GGA ACC AGT GGG CAA CTC TGG TAT CAT GGG GAA ATA Glu Leu Pro Gln Gly Thr Ser Gly Gln Leu Trp Tyr His Gly Glu Ile 305 310 315			961
CTA TTT ATG CGA CAT GGG AAA AGT CTA ACT GTG ATT CCA TAC AAG TGC Leu Phe Met Arg His Gly Lys Ser Leu Thr Val Ile Pro Tyr Lys Cys 320 325 330			1009
GAA GCA TCA TCT CTG GCG AGC GCT CTT GGG AAG CTC AAG CAT ACT CAA Glu Ala Ser Ser Leu Ala Ser Ala Leu Gly Lys Leu Lys His Thr Gln 335 340 345			1057
GAG TCA GGC AGT CAT TCC GTG CCC CAT TTT GTA AAC TGG GAG ACG TGT			1105

Glu 350	Ser	Gly	Ser	His	Ser 355	Val	Pro	His	Phe	Val 360	Asn	Trp	Glu	Thr	Cys 365	
TCA Ser	GAA Glu	TAT Tyr	GAA Glu	CTT Leu 370	GGG Gly	TCC Ser	TAC Tyr	AGT Ser	GCA Ala 375	AAG Lys	CAG Gln	ACA Thr	AGA Arg	ACT Thr 380	CTT Leu	1153
AGA Arg	AAA Lys	AAA Lys	ACT Thr 385	GAA Glu	ACA Thr	AAT Asn	TTA Leu	CAC His 390	CCA Pro	GAA Glu	GTT Val	CCA Pro	GGA Gly 395	GTC Val	AAA Lys	1201
GAA Glu	CTT Leu	TTA Leu 400	TCA Ser	ATA Ile	ATA Ile	AAG Lys	AAA Lys 405	GAT Asp	TCA Ser	GAA Glu	AAG Lys	CAC His 410	ATT Ile	GAA Glu	GTA Val	1249
GAA Glu 415	CTA Leu	CGT Arg	AAG Lys	TTT Phe	TTG Leu	GCT Ala 420	AAG Lys	TCG Ser	ACA Thr	CCT Pro	GAC Asp 425	TTT Phe	CAT His	ACT Thr	ATA Ile	1297
ATT Ile 430	GGA Gly	GAC Asp	ATA Ile	GTG Val	TCT Ser 435	GGA Gly	CTT Leu	CTG Leu	GGA Gly	AGA Arg 440	TGT Cys	AAA Lys	GTG Val	GAA Glu	CCA Pro 445	1345
TCG Ser	TTT Phe	TAC Tyr	CCC Pro	CGG Arg 450	AAC Asn	TGT Cys	CTG Leu	ATG Met	CAG Gln 455	CTC Leu	ATC Ile	CAG Gln	ACA Thr	CAC His 460	GTG Val	1393
CTT Leu	TCC Ser	TAC Tyr	AGC Ser 465	TTA Leu	TGC Cys	CCT Pro	GAC Asp	TTG Leu 470	ATG Met	GAG Glu	ATT Ile	GCC Ala	CTA Leu	GAG Glu	CAC His	1441
ACA Thr	GAT Asp	GTG Val 480	CAG Gln	ATG Met	TTA Leu	CAG Gln	CTG Leu 485	TGT Cys	CTG Leu	CAG Gln	CAG Gln	TTC Phe 490	CCT Pro	GAC Asp	ATT Ile	1489
CCT Pro 495	GAG Glu	TCC Ser	ACC Thr	ACC Thr	TGT Cys	GCT Ala 500	TGC Cys	TTA Leu	AAA Lys	CTT Leu	TTC Phe 505	CTG Leu	AGC Ser	ATT Ile	GGT Gly	1537
GAT Asp 510	GAC Asp	TGT Cys	CTT Leu	CGG Arg 515	GAC Asp	AGT Ser	AAT Asn	GTC Val	AGC Ser 520	ATG Met	GAG Glu	TCA Ser	GTT Val	TTT Phe	GAT Asp 525	1585
TAT Tyr	GGT Gly	GAT Asp	ACT Thr 530	ACA Thr	CAA Gln	GAT Asp	GAG Glu	AAA Lys	AAG Lys 535	GAA Glu	ATG Met	GAA Glu	GAG Glu	CAA Gln 540	ACT Thr	1633
GAA Glu	ATT Ile	GTT Val	CAG Gln 545	AAT Asn	GGC Gly	TTC Phe	GGT Gly	CCT Pro 550	GAA Glu	GAT Asp	GGT Gly	GGC Gly	TGC Cys 555	AGT Ser	AAA Lys	1681
GAT Asp	AGT Ser	CAG Gln 560	CAG Gln	TTG Leu	AAC Asn	AAA Lys	AAG Lys 565	CCT Pro	GGT Gly	GAC Asp	ACA Thr	GCA Ala 570	CAG Gln	GAG Glu	CCC Pro	1729
ATC Ile 575	TCC Ser	TTC Phe	CCT Pro	GTG Val	ACC Thr	TTG Leu 580	TGT Cys	CCT Pro	GTG Val	GCA Ala	CCA Pro	AAG Lys	CGA Arg	GCA Ala	GCT Ala	1777
CTG Leu 590	CTA Leu	AAT Asn	GCA Ala	GTC Val	CTT Leu 595	CAT His	TCA Ser	GCA Ala	TAC Tyr	AGT Ser 600	GAG Glu	CCC Pro	TTC Phe	CTC Leu	CTG Leu 605	1825

CCG CAC TTG AAG GAC ATC CCT GCG AAG CAT ATA ACG CTG TTT CTG CAG Pro His Leu Lys Asp Ile Pro Ala Lys His Ile Thr Leu Phe Leu Gln 610 615 620	1873
TAT TTG TAT TTC CTC TAT TTG AAG TGC ACT GGC AGT GCT ACC ATG ACT Tyr Leu Tyr Phe Leu Tyr Leu Lys Cys Thr Gly Ser Ala Thr Met Thr 625 630 635	1921
CTC CCT GGA GTA AGC CCT CCA ACC GTG AGC CAG ATC ATG GAT TGG ATA Leu Pro Gly Val Ser Pro Pro Thr Val Ser Gln Ile Met Asp Trp Ile 640 645 650	1969
TGC CTA CTT CTA GAT GCT AAT TTT ACT GTC GTC TTA ATG ATA CCA GAA Cys Leu Leu Leu Asp Ala Asn Phe Thr Val Val Leu Met Ile Pro Glu 655 660 665	2017
GCA AAA AGA CTT TTG CTT AAT CTT TAC AAT TTT GTG AAA TCG CAG ATC Ala Lys Arg Leu Leu Leu Asn Leu Tyr Asn Phe Val Lys Ser Gln Ile 670 675 680 685	2065
TCC ATC TAC TCC GAG CTC AAC AAG ATT GAA GTC AGC TTC CGG GAG CTG Ser Ile Tyr Ser Glu Leu Asn Lys Ile Glu Val Ser Phe Arg Glu Leu 690 695 700	2113
CAG AGG TTA AAT CAG GAG AAG AGC AGT AGA GGA CTG TAC TCC ATC GAA Gln Arg Leu Asn Gln Glu Lys Ser Ser Arg Gly Leu Tyr Ser Ile Glu 705 710 715	2161
GTT CTG GAA CTT TTC TGACAGCCAG AGACAGTCTT ACTGGTTGTC CTAGAACTGG Val Leu Glu Leu Phe 720	2216
ATATATAGAC CAAGCTGGCC TCAAACTCAC AGAGATCCAA CCCAGAGCAC TGCGATTAAA	2276
GGTGTGGCCA CCACACCTGG CTTAATGGAT TTTTATATAT GAATGTACCT GTTCATCCAG	2336
GGGAGAAAGT TTTGTTTGTA TCCATCTCAG TGTCCAGAGA GGAGTGTGAG TCTCTGCACC	2396
ATCACATACC ACCATTCCCTC AGTGCTACGG GTTCCACAGA CATCACTTCT GGGTTGCAGA	2456
CATCACTGGA AGAATGTTAC AAACCTTTTAA AACATGGAAT TGAGTTGATT TTAAGTAAAC	2516
TTATTTGTGT ACTGATAAAA A	2537

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 722 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Ala Ala Leu Glu Glu Glu Phe Thr Leu Ser Thr Gly Val Leu Gly  
1 5 10 15

Ala Gly Pro Glu Gly Phe Leu Gly Val Glu Pro Ser Asp Lys Ala Asp  
20 25 30

Gln Phe Leu Val Thr Asp Arg Gly Arg Thr Val Val Leu Tyr Lys Val



35					40					45					
Ser	Asp	Gln	Lys	Pro	Leu	Gly	Ser	Trp	Ser	Val	Lys	Gln	Gly	Gln	Thr
50						55					60				
Ile	Thr	Cys	Pro	Ala	Val	Cys	Asn	Phe	Gln	Thr	Gly	Glu	Tyr	Ile	Met
65						70					75				80
Val	His	Asp	His	Lys	Val	Leu	Arg	Ile	Trp	Asn	Asn	Glu	Asp	Val	Asn
				85					90					95	
Leu	Asp	Lys	Val	Phe	Lys	Ala	Thr	Leu	Ser	Ala	Glu	Val	His	Arg	Ile
			100					105					110		
His	Ser	Val	Gln	Arg	Thr	Glu	Pro	Leu	Val	Leu	Phe	Arg	Arg	Gly	Ala
		115					120					125			
Ala	Arg	Gly	Leu	Glu	Ala	Leu	Leu	Ala	Glu	Pro	Gln	Gln	Asn	Ile	Glu
		130				135					140				
Ala	Val	Ile	Pro	Asp	Glu	Glu	Val	Ile	Glu	Trp	Ser	Glu	Val	Phe	Met
145						150					155				160
Leu	Phe	Lys	Gln	Pro	Val	Leu	Ile	Phe	Ile	Thr	Glu	Asn	His	Gly	Ser
				165					170					175	
Tyr	Val	Ala	Tyr	Val	Gln	Leu	Cys	Lys	Ser	His	Ser	Leu	Ser	Lys	Tyr
			180					185					190		
Ile	Leu	Leu	Leu	Gly	Lys	Glu	Glu	Lys	Ser	Ala	Lys	Pro	Asn	Phe	Thr
		195					200					205			
Ala	Arg	Val	Asp	Gly	Lys	Phe	Ile	Ser	Leu	Val	Ser	Leu	Gly	Ser	Asp
		210				215					220				
Gly	Cys	Val	Tyr	Gly	Thr	Leu	Val	Pro	Ile	Tyr	Ser	Ser	Asp	Thr	Glu
225						230					235				240
Asn	Asn	Gln	Arg	Leu	Val	Arg	Ala	Leu	Met	Leu	Lys	Thr	Val	Val	Ser
				245					250					255	
Gly	Ser	Ala	Arg	Asn	Gly	Ser	Ala	Leu	Thr	Ile	Leu	Asp	Gln	Asp	His
			260					265					270		
Ile	Ala	Val	Leu	Gly	Pro	Pro	Leu	Pro	Ala	Ser	Lys	Glu	Cys	Leu	Ser
		275					280					285			
Ile	Trp	Asn	Ile	Lys	Phe	Gln	Thr	Leu	Gln	Thr	Ser	Lys	Glu	Leu	Pro
		290				295					300				
Gln	Gly	Thr	Ser	Gly	Gln	Leu	Trp	Tyr	His	Gly	Glu	Ile	Leu	Phe	Met
305						310					315				320
Arg	His	Gly	Lys	Ser	Leu	Thr	Val	Ile	Pro	Tyr	Lys	Cys	Glu	Ala	Ser
				325					330					335	
Ser	Leu	Ala	Ser	Ala	Leu	Gly	Lys	Leu	Lys	His	Thr	Gln	Glu	Ser	Gly
			340					345				350			
Ser	His	Ser	Val	Pro	His	Phe	Val	Asn	Trp	Glu	Thr	Cys	Ser	Glu	Tyr
		355					360					365			
Glu	Leu	Gly	Ser	Tyr	Ser	Ala	Lys	Gln	Thr	Arg	Thr	Leu	Arg	Lys	Lys

370	375	380
Thr Glu Thr Asn Leu His 385	Pro Glu Val 390	Pro Gly Val Lys Glu Leu Leu 395 400
Ser Ile Ile Lys 405	Lys Asp Ser Glu Lys 410	His Ile Glu Val Glu Leu Arg 415
Lys Phe Leu Ala Lys Ser Thr Pro Asp Phe His Thr Ile Ile Gly Asp 420 425 430		
Ile Val Ser Gly Leu Leu Gly Arg Cys Lys Val Glu Pro Ser Phe Tyr 435 440 445		
Pro Arg Asn Cys Leu Met Gln Leu Ile Gln Thr His Val Leu Ser Tyr 450 455 460		
Ser Leu Cys Pro Asp Leu Met Glu Ile Ala Leu Glu His Thr Asp Val 465 470 475 480		
Gln Met Leu Gln Leu Cys Leu Gln Gln Phe Pro Asp Ile Pro Glu Ser 485 490 495		
Thr Thr Cys Ala Cys Leu Lys Leu Phe Leu Ser Ile Gly Asp Asp Cys 500 505 510		
Leu Arg Asp Ser Asn Val Ser Met Glu Ser Val Phe Asp Tyr Gly Asp 515 520 525		
Thr Thr Gln Asp Glu Lys Lys Glu Met Glu Glu Gln Thr Glu Ile Val 530 535 540		
Gln Asn Gly Phe Gly Pro Glu Asp Gly Gly Cys Ser Lys Asp Ser Gln 545 550 555 560		
Gln Leu Asn Lys Lys Pro Gly Asp Thr Ala Gln Glu Pro Ile Ser Phe 565 570 575		
Pro Val Thr Leu Cys Pro Val Ala Pro Lys Arg Ala Ala Leu Leu Asn 580 585 590		
Ala Val Leu His Ser Ala Tyr Ser Glu Pro Phe Leu Leu Pro His Leu 595 600 605		
Lys Asp Ile Pro Ala Lys His Ile Thr Leu Phe Leu Gln Tyr Leu Tyr 610 615 620		
Phe Leu Tyr Leu Lys Cys Thr Gly Ser Ala Thr Met Thr Leu Pro Gly 625 630 635 640		
Val Ser Pro Pro Thr Val Ser Gln Ile Met Asp Trp Ile Cys Leu Leu 645 650 655		
Leu Asp Ala Asn Phe Thr Val Val Leu Met Ile Pro Glu Ala Lys Arg 660 665 670		
Leu Leu Leu Asn Leu Tyr Asn Phe Val Lys Ser Gln Ile Ser Ile Tyr 675 680 685		
Ser Glu Leu Asn Lys Ile Glu Val Ser Phe Arg Glu Leu Gln Arg Leu 690 695 700		
Asn Gln Glu Lys Ser Ser Arg Gly Leu Tyr Ser Ile Glu Val Leu Glu		

705

710

715

720

Leu Phe

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW016

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 385..748
- (D) OTHER INFORMATION: /label= SAC\_24530

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATTATTCTCC CCTTGTCATT AAAAGAAATC AGCAAAACCA AACAACTGGC TACACCACGA	60
ATTGTCGTTA AATTTGCTAA CTGGTGTCTA AAAGCCGTGT AGCTACCTCG GTCCTGCTTG	120
CTAGGTTTGC CACTAGAAGG AAGCATACTT AAAACAATGG CTACTTGGAT CCTCAGGGAG	180
ATCCTGTCTG CAGTCGCGTG GTCACCCCTTA GCTTCATCAA AGCACTAACA GCTCACCCGG	240
CCAGGCTTCA TGAGCACTGA CCCTCAAGCA AGCAGGTTTA TTAAACATTT AGATGCCAAC	300
CTCACTTACT GTTTCCTGCA GTCATGGAGA GTTTACTTAA CAAGTTTGTA AATAATAAAC	360
TGGCACTTTG CACACAGACT TGGTACTATC CTAGGGGAAG GCCTGCTTTA TTTGGTTTCT	420
AGACCGAGTA GGAAGTGATC CATTTACCAC TGAGGGCAGC CCCATTCAGA GTCTTAAGTG	480
ACTAAGCCAG TGTTGAACAA GCAATTTCCA GGCTTTGTTC TTCAGGGAAC TTCCCATCAG	540
CTTTGAAGTC GGTCCGTGTC ACCCTAGGCA CATGGATCAG TTCACAAGTG GGGTTCAGTG	600
GAGAGAACTT CCCCCTCAGA AGTCACTTGA AACTTAGATG AGATTTGGGA CACTTGCTGG	660
TTGACTCTGT CTCATTTGTG TAAAAAGTAG TTTNTNTNAA NANTNGGTTT TTTTTTTTTT	720
TTCAAGGTTA TACTTTGTCC CATTCCTANG TTANTACAAA GTCTTGAAAG GGCCTTTGTA	780
GGGCTTTTTA ANNCAGGGTC TTAACATATG AACTCTGGCT TGGCCTGGAA CTTGCTATGT	840
AGACCAGGTT ACCCTCAAAC TTGCCTGTCT TCCCAAATAC TGGGATTAAG GTTTCTGTGA	900
CCATACCTGG CTTTACCTGA TGAATTCCTA AACACCAGAA AACCAGTACT GTATGAGATG	960

```

TTAATGTGTG TTCCTTTCAG ACTGGAGTAC AGACCAGTAG ATAACAGATA ACAGCTGGTT      1020
CACCTTAATC TGCCTTTTTG TGTATTAATC TGTGTTTAGA GAACGGAACA ATAGCCAGAA      1080
TTNNNCTAGC GAGTTCGAGG CCAGTTGGTG TATATGTGGG ACTCTTAACC AAAACAGCAA      1140
GCGTTCCCTG GGGTAGTTCA CAATGATCTC CAGCTTCCTT GTTAACCAGA TAACTGCNAN      1200
TCCAGATGTA TGACCCTGGT TGGTTTATTG TATTGATATG TTTCTGTAAT ATGAGTAAAT      1260
TATTGTTACT TAAAAGTAAT AAATTCAAGT TTCAAAATTG AAAAAA                      1305

```

## (2) INFORMATION FOR SEQ ID NO:12;

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 955 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW017

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 393..627
- (D) OTHER INFORMATION: /label= SAC\_24091

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

GTCGACAATC CTGTTATGAC AAGCCCGTCT CGATACTGAC AACTGAAGTG GAAATTGCTT      60
GAACATCAAG ATAGCTAACT CGGAAAGAAT CCCAGACTTT GGATAGTTTC TGAGTCTTCT      120
AGAATTTTCC AAGAAGAAAC CATGGTGGCA AATGCGGAGA CTCGGGCGGG CACTGGGAAC      180
ACAGCACAGT GGTCTTAGGG AGGTGCTTTG TCAGGAATGA ACAGTCATGG TTATAATCCA      240
CGTTTCCATT GCTACTCACG AATGATTCTC TTCTGTTTTG TTTTAAAAT TTTTACAC      300
TGAATTTCTA TTTAGACACT AAAACATATA GGGGTGCTTG TCACCCGGAT ACATTTATCT      360
GTGAGCCAGC TATTCGGATG TCATGGCTGG GTACCTAACT TACTTCATA TGTGAAGTGT      420
GCTAAACTCA AACCAGTTTA CAGAAACGAT GTATTTTGTG TATAGTAAAT TGTATATTCA      480
CCCTTTTACC ACGGCCGGTT TTTTAAACAA ATGAATACTC TAGATTTTTC TTCCAAATGA      540
GGTTACTGTT GGGGTGGGGT TGAAGTAGTG ATGCTGTAGA AGGGAATCCG CATGCACTAA      600
AAAGTGTGTC TGCCTAAAAG TGTGTACAGC AGGCACAACC TTCCTTGTGG ATTTCCGTCT      660
GCTCACTGCA GATCTGCCTG TGGTTTAGAA ATAGAATTCA AGAGCCATCA AGGAGTGACC      720
AGCTTGACAC CACTGCCAAA TTCAGAATGA GGAACCTTGA GAGAAGGAAC TGTCGCTCAG      780
CCAGAAGCAG ATGAAGCCAC TGGTCCTGGT TCACAGTTTA GTGTCATGCT CTGGTTTATA      840

```

(2) INFORMATION FOR SEQ ID NO:13:

(A) LENGTH: 1198 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(B) CLONE: HW018

(A) NAME/KEY: CDS  
(B) LOCATION: 86..1102  
(D) OTHER INFORMATION: /product= "predicted polypeptide"

(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (644..994)  
(D) OTHER INFORMATION: /label= SAC\_23880

GTCGACCTGG	GAGTCCGGGC	AGCGAGCCCC	TCGCCGTCGG	GTCCCACGAA	GAAACTGGTG	60										
GCTGAAGCCA	TATTCTGCAG	CATTC	ATG	GAG	ATC	AGT	AAA	AAT	ATT	ACG	AAC	112				
			Met	Glu	Ile	Ser	Lys	Asn	Ile	Thr	Asn					
			1				5									
CCA	GAA	GAA	CTT	TGG	AAA	ATG	AAG	CCT	AAG	GGG	AAC	CTG	GAA	GAT	GAC	160
Pro	Glu	Glu	Leu	Trp	Lys	Met	Lys	Pro	Lys	Gly	Asn	Leu	Glu	Asp	Asp	
10					15					20					25	
AGT	TAC	TCG	ACT	AAG	GAC	ACC	GGA	GAG	ACA	AGC	ATG	CTG	AAA	AGA	CCG	208
Ser	Tyr	Ser	Thr	Lys	Asp	Thr	Gly	Glu	Thr	Ser	Met	Leu	Lys	Arg	Pro	
				30					35					40		
GTG	CTC	TCG	CCT	TTG	CCG	CAC	ACG	GTC	CAC	GTC	GAC	GCC	TTC	GAC	TGC	256
Val	Leu	Ser	Pro	Leu	Pro	His	Thr	Val	His	Val	Asp	Ala	Phe	Asp	Cys	
			45					50					55			
CCC	ACG	GAG	CTC	CAG	CAC	ACG	CAG	GAA	CTC	TTT	CCA	AAC	TGG	CAA	TTG	304
Pro	Thr	Glu	Leu	Gln	His	Thr	Gln	Glu	Leu	Phe	Pro	Asn	Trp	Gln	Leu	
		60					65				70					
CCA	ATT	AAA	GTT	GCC	GCC	GTC	CTA	TCG	TCT	CTC	ACC	TTC	CTG	TAC	ACT	352
Pro	Ile	Lys	Val	Ala	Ala	Val	Leu	Ser	Ser	Leu	Thr	Phe	Leu	Tyr	Thr	
	75					80					85					
CTT	CTG	AGG	GAA	ATC	ATT	TAC	CCA	TTA	GTC	GCT	TCC	CGC	GAA	CAG	TAT	400
Leu	Leu	Arg	Glu	Ile	Ile	Tyr	Pro	Leu	Val	Ala	Ser	Arg	Glu	Gln	Tyr	
90					95					100					105	

TTT	TAT	AAG	ATC	CCG	ATC	CTG	GTC	GTT	AAC	AAA	GTC	TTG	CCA	ATG	GTC	448
Phe	Tyr	Lys	Ile	Pro	Ile	Leu	Val	Val	Asn	Lys	Val	Leu	Pro	Met	Val	
				110					115					120		
TCC	ATT	ACC	CTC	TTG	GCA	TTG	GTT	TAT	CTG	CCA	GGA	GAG	ATA	GCG	GCA	496
Ser	Ile	Thr	Leu	Leu	Ala	Leu	Val	Tyr	Leu	Pro	Gly	Glu	Ile	Ala	Ala	
			125					130					135			
GTT	GTG	CAG	CTT	CGC	AAT	GGG	ACC	AAG	TAC	AAG	AAG	TTC	CCA	CCC	TGG	544
Val	Val	Gln	Leu	Arg	Asn	Gly	Thr	Lys	Tyr	Lys	Lys	Phe	Pro	Pro	Trp	
		140					145					150				
TTA	GAC	AGA	TGG	ATG	CTG	GCG	AGG	AAA	CAA	TTT	GGG	CTC	CTC	AGC	TTC	592
Leu	Asp	Arg	Trp	Met	Leu	Ala	Arg	Lys	Gln	Phe	Gly	Leu	Leu	Ser	Phe	
	155					160					165					
TTC	TTT	GCA	GTT	CTG	CAC	GCC	ATT	TAC	AGT	CTC	TCG	TAC	CCA	ATG	AGA	640
Phe	Phe	Ala	Val	Leu	His	Ala	Ile	Tyr	Ser	Leu	Ser	Tyr	Pro	Met	Arg	
170					175					180					185	
CGG	TCC	TAC	AGA	TAC	AAG	CTG	CTC	AAC	TGG	GCT	TAC	AAA	CAG	GTT	CAA	688
Arg	Ser	Tyr	Arg	Tyr	Lys	Leu	Leu	Asn	Trp	Ala	Tyr	Lys	Gln	Val	Gln	
				190					195					200		
CAA	AGC	AAA	GAA	GAT	GCC	TGG	GTT	GAG	CAT	GAT	GTC	TGG	AGA	ATG	GAG	736
Gln	Ser	Lys	Glu	Asp	Ala	Trp	Val	Glu	His	Asp	Val	Trp	Arg	Met	Glu	
		205						210					215			
ATT	TAT	GTG	TCC	CTG	GGG	ATT	GTG	GGA	CTG	GCC	ATC	TTG	GCT	CTC	TTG	784
Ile	Tyr	Val	Ser	Leu	Gly	Ile	Val	Gly	Leu	Ala	Ile	Leu	Ala	Leu	Leu	
		220					225					230				
GCT	GTG	ACA	TCT	ATC	CCA	TCT	GTG	AGC	GAC	TCT	TTA	ACC	TGG	AGA	GAA	832
Ala	Val	Thr	Ser	Ile	Pro	Ser	Val	Ser	Asp	Ser	Leu	Thr	Trp	Arg	Glu	
	235					240					245					
TTT	CAC	TAT	ATT	CAG	AGC	AAA	CTG	GGA	ATT	GTC	TCT	CTT	CTT	CTG	GGC	880
Phe	His	Tyr	Ile	Gln	Ser	Lys	Leu	Gly	Ile	Val	Ser	Leu	Leu	Leu	Gly	
250					255					260					265	
ACG	GTA	CAC	GCT	TCG	ATT	TTT	GCC	TGG	AAT	AAA	TGG	GTA	GAT	ATC	AGT	928
Thr	Val	His	Ala	Ser	Ile	Phe	Ala	Trp	Asn	Lys	Trp	Val	Asp	Ile	Ser	
				270					275					280		
CAG	TTT	GTC	TGG	TAC	ATG	CCT	CCG	ACT	TTC	ATG	ATA	GCT	GTT	TTC	CTT	976
Gln	Phe	Val	Trp	Tyr	Met	Pro	Pro	Thr	Phe	Met	Ile	Ala	Val	Phe	Leu	
			285					290					295			
CCA	ACT	GTT	GTT	CTG	ATC	TGT	AAA	ATT	GTA	CTT	TGC	CTG	CCC	TGC	CTG	1024
Pro	Thr	Val	Val	Leu	Ile	Cys	Lys	Ile	Val	Leu	Cys	Leu	Pro	Cys	Leu	
		300					305					310				
AGG	AAG	AAG	ATA	CTG	AAG	ATT	AGA	TGT	GGT	TGG	GAA	GAT	GTC	CGC	AAA	1072
Arg	Lys	Lys	Ile	Leu	Lys	Ile	Arg	Cys	Gly	Trp	Glu	Asp	Val	Arg	Lys	
	315					320					325					
GTT	AAC	AGG	ACT	GAG	ATG	GCC	TGC	AGG	TTG	TAGAACTACT	GTTTGCACAT					1122
Val	Asn	Arg	Thr	Glu	Met	Ala	Cys	Arg	Leu							
330					335											
AATTGCTTAA	TACCGATGTC	TTATAAACAT	TTCAAACCTG	TGTCTGTAA	TAAAATGAAT											1182

ACTTCTGGAT CAAAAA

1198

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```

Met Glu Ile Ser Lys Asn Ile Thr Asn Pro Glu Glu Leu Trp Lys Met
 1           5           10           15
Lys Pro Lys Gly Asn Leu Glu Asp Asp Ser Tyr Ser Thr Lys Asp Thr
          20           25           30
Gly Glu Thr Ser Met Leu Lys Arg Pro Val Leu Ser Pro Leu Pro His
          35           40           45
Thr Val His Val Asp Ala Phe Asp Cys Pro Thr Glu Leu Gln His Thr
          50           55           60
Gln Glu Leu Phe Pro Asn Trp Gln Leu Pro Ile Lys Val Ala Ala Val
          65           70           75           80
Leu Ser Ser Leu Thr Phe Leu Tyr Thr Leu Leu Arg Glu Ile Ile Tyr
          85           90           95
Pro Leu Val Ala Ser Arg Glu Gln Tyr Phe Tyr Lys Ile Pro Ile Leu
          100          105          110
Val Val Asn Lys Val Leu Pro Met Val Ser Ile Thr Leu Leu Ala Leu
          115          120          125
Val Tyr Leu Pro Gly Glu Ile Ala Ala Val Val Gln Leu Arg Asn Gly
          130          135          140
Thr Lys Tyr Lys Lys Phe Pro Pro Trp Leu Asp Arg Trp Met Leu Ala
          145          150          155          160
Arg Lys Gln Phe Gly Leu Leu Ser Phe Phe Phe Ala Val Leu His Ala
          165          170          175
Ile Tyr Ser Leu Ser Tyr Pro Met Arg Arg Ser Tyr Arg Tyr Lys Leu
          180          185          190
Leu Asn Trp Ala Tyr Lys Gln Val Gln Gln Ser Lys Glu Asp Ala Trp
          195          200          205
Val Glu His Asp Val Trp Arg Met Glu Ile Tyr Val Ser Leu Gly Ile
          210          215          220
Val Gly Leu Ala Ile Leu Ala Leu Leu Ala Val Thr Ser Ile Pro Ser
          225          230          235          240
Val Ser Asp Ser Leu Thr Trp Arg Glu Phe His Tyr Ile Gln Ser Lys
          245          250          255
Leu Gly Ile Val Ser Leu Leu Leu Gly Thr Val His Ala Ser Ile Phe

```

260	265	270
Ala Trp Asn Lys Trp Val Asp Ile Ser Gln Phe Val Trp Tyr Met Pro		
275	280	285
Pro Thr Phe Met Ile Ala Val Phe Leu Pro Thr Val Val Leu Ile Cys		
290	295	300
Lys Ile Val Leu Cys Leu Pro Cys Leu Arg Lys Lys Ile Leu Lys Ile		
305	310	315
Arg Cys Gly Trp Glu Asp Val Arg Lys Val Asn Arg Thr Glu Met Ala		
325	330	335
Cys Arg Leu		

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1801 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW032

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1412..1793)
- (D) OTHER INFORMATION: /label= SAC\_24831

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TTTTTCCAGC TTTTAAAGTT GCATTATTT TTGATAATCT GTACTGTAGC CAATCAATTC	60
TATTCTCCAT TCTTGCTTAA TTGTATTAGC CATAAAGGAG TTCAGGTTCA TAGGACTCAG	120
TTTCTCTATT AAACCGTCAT TGATTATGGA AATACTGTCT TAAGCACTAT GCAGATTAAC	180
ACCATACTTT TCTTCCATTA CATCTTCAA GACAGTCACT CACAACCTAG AAGACTGATA	240
CAGCATTAGA TGAACAAAAC AAGATTCACA GATATATGTA GTGAAACTCA AACTAGGAGT	300
CTTAGAGATA AAAAAGATAA GTAGATATTT AAAATGTAGT CTGTAAAGA TTTAATCGGT	360
TTAAAGTATA AGATTAGAAA GACAAAGTTT CTATTGTTTT GAAAGTTTAA AGTAAAGAAA	420
AAAAAACTGG AAAATATGTC TTTGTTGTTT TTTCTCCGTT CTTAATTACG AGCTATCCAA	480
CAGAAACCCC TAGGTCATAG TTTATGTTTC TGTTCATC AATGTAAC TAACACTGGA	540
GAAAGAAAAA GGAAGACATA TTAACCAATG GCCGCAATCA CTTTTCAC TTGCCACGGT	600
ACAAATCTGA TTTGGATGAA ATGGTTTGTG GCATAAAAT TTTAAATTTT AGAAGCCAAT	660



GTGGTAGATT TCTCCATGTT TGCACAAATC CAATTTCAAC TACTTTTACT TCACTAATTA	720
ATTAAATGGA TTGGAAACAT TCTAGAAATG AAATTGTGTC TTTAGATGGA GTTGGGGAGA	780
GTACGCTTTA TTTTAAATAT TTAGTTCAAA TTTCATTTTC ACTTCTTGGT CTTATTATTT	840
CAGCTTCAGT TGGGGCTGGC TTCTGTGGGG CTTACATGCT GACTGTGTCA GGCAGCTGGA	900
GGTGATCTAC ATTGACAGGA TTCAATGACT TTCTCCTCCG CGAAAAGCCC CTGTCAGCTA	960
CACAAGGATG GGAGGTTTTT GGCACACTGT CAAGTGATTG CAGATCCTGA CTTCTCAAAT	1020
ATGAGGTGTC GGCAGTGGCA GCAGCAGCTA CTGGTAGCTG CCCCAGACTC CTCAGATAAC	1080
TTACACAGGA GCTGGTTTTAT GCATCCCCTT CGTCAGAGCT TCCTCTGGGC AAGTGAAAGT	1140
TGCAAGCCTT CACTTGCCCC CATGTCTCAG CTCTCTAAGT GACCCTGCCT GTGCTGGCCC	1200
TTGTGCTTAT GGTGGTGGTG GTGGTGGTGC AAGCCTGAAG GAGGAAGACT TGTCTCAACA	1260
TCTAATGCCC CTGGTTGCTG ATTCTCTGAA GGCTTACTGC TCCCAGCAAC CGTCAGCCTC	1320
AGTCGCAGCC GGGCCTCGCT CCTCAACTTG GCAAAAATGC CTACAGAGAC TGAGAGATGC	1380
ATCGAGTCCC TGATTGCTGT TTTCCAGAAG TACAGTGGGA AGGATGGAAA TAGCTGTCAT	1440
CTCTCCAAAA CTGAGTTCTT TTCCTTCATG AACACGGAGC TGGCCGCCTT CACGAAGAAC	1500
CAGAAGGACC CCGGTGTCCT CGACCGCATG ATGAAGAAGC TGGACCTCAA CAGTGATGGG	1560
CAGCTAGATT TCCAAGAGTT TCTCAACCTT ATTGGTGGCT TAGCTATAGC ATGCCATGAG	1620
TCCTTCCTCC AGACTTCCCA GAAGCGTATC TAACCCTCTC CATTCCCTTC CAGCCACCAA	1680
GTCATCGCCT CCTCCACTCC TTCCCCCATC CACACCTGCA CTGAGCCCAC CACACCTACC	1740
ACACATGCAG CCCACGCCTG ACAGGGAAAA TAAACAATG TCATTTTTTT AAATGTAAAA	1800
A	1801

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2646 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW033

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 50..263
- (D) OTHER INFORMATION: /label= SAC\_23901

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCGACTTGA CTTCTTAGTT TTCTGACGGG AGCTCTCTTC CTTCTGAGTA CTAGTCTCGG	60
GCCTCCAGCC CCAGGAAGAC CCTCTGCTTT TCTCTCCACC CTTGGCTTGC TCCTCAGCAG	120
GGAAGAAGTA GATCACAGTC TCACTGCCTT CAGTTCCTCAG CTTTGCCAGT CACCAGCTGT	180
GTAGTCTCAG GCCAATCACC ACCTCCTCTA GCCTAGGTAG TTGGTTTAT GTGGGTGAAT	240
GTTTTGCCTG GATGTATGTA TGTACACCAC CTGTGTGCCT GGTGCCCCCTG CAGTCAGAAG	300
AAAGTGCCAC ACCCCTAGGA ACTGGAGTTA AGATGGGTTA GAGCCTCCAT GTGGGTGCTG	360
AGAACTGAAC CTGGTTCTCT GGAAAAGCAG CCGGCACTCA ACCTCTGAGC CAGCTCTTCA	420
TCCCCGAGTC CGTTTATTTA ACCGGCTATA AAATGGATT AAAACGACTT ACCTCACCAG	480
GCCTGTGAAA CATTTAAAAA CTTAATAAAA ATTCCTACACT TAATTGTTTT TGCTTATGCT	540
TTCTAGTCCC TTTAGTCTTA GACAGTTTCT GTTTCTTGTA TTTATAAAGT TGGTATTAAA	600
TATATAGCCA GTGGACAAGA ACCTTACTGC TATGGGCTGG CCCTGTAGG GTCCTAGACT	660
AGAGGAGCGT ACTTTGTGGT TTAAAGTGGT TTGGCCTGCC TACCTGTTTT TCCCACTGGA	720
CTCAAAACAC TAATAAAGCA GGGGTGACCT GTGTCACCTG AGCATTGAGC AGAACACTGG	780
CCCAGGCAAT GGCCTCAAAG CTGGGCTACA TTATAAAGTG AACTTTCTTG TTCATACTGG	840
GTCACAGAAG CAGGCCCAGA AAAAATAAAA AATCCTGGTC ACTCGGTTCC CCCTGTCACC	900
ATCATCTGTT CACTAGTGCC CCTCTGACCA TTCTCACCCA AGACAGGCAT CCTCTGGGAA	960
CCTTAGGCAG GTCCATTTCT TCTTTGTCAT TTGGAATGAA GCACTAACAG CCCTGGGGTC	1020
TCTGCCATTG GCTTCCCAGA GTTGTCCAAT GAGTTTTCGT GAAGGAATCT GGGTTTATCT	1080
TTCCCCGGCT TCCTGCCCTG TTAGCAGAAT TTCAAATAGC AGATAATAGA TTCTGACCTG	1140
TGAGGCCATT TTTCTCTCAA TCCCTTTCTT CTGAGACATT AGCCTTGGGC ATGCAGCAGC	1200
TGCGCACTGG TTTCAGGGAG GAAAGGGGTG AATGGTGAGC GAACCTGGCC CTGTGGATGC	1260
TAACCTCACT GAGATGAAAA CATGGGGGCT TTTCTCTGAA GACGCAGGGC CGTGTAACA	1320
TATGACTTTG ACCATATTCT CAGGTTTGGG GATGTGGGAA GTTCCAGGAA CAGAAATCCC	1380
CTGCTGAGTT CTTGGGCACT GTGGCTATTT GTTCCCATGA GGACCTCTCC CTGCAGTCTG	1440
CCCTCTGGCA GGGTGGAGTA TAGAAGACTC AAGAGCTACA GATGCCCCCG TCACTGGAGC	1500
TCCCCTGCAC CTAGGCCATC AATCACGTCC ACCATCACCA GAGGGGCTTA GGCCTAGGGT	1560
GTATGCTGAG GGAAAGACGA GGGCCTGTCC CACTACTCGA GTAGAGCAGA TGGCCTTCTG	1620
TGTAGGCACT GAATTTTATA GCAGTGCCTT GAGGAGCCCA GCATGTTCTC AAGCACCTCC	1680
TCCCAGTTTA TATTTGGAAG TCTGCTTCAC AGGGACCAGA GGGCATAAAC CTAACCTACC	1740
CACTTCCCAT CAGGGACAGC CCAAAGCACA CCAATATATG TAAGGGGTGC AGGGAGGTAT	1800
CATGGGTCTC TTCTCTCCTC TCAGTGCTCA AACTCAGGTG GAGGCTATGG ATTCAGGATC	1860

```

TTGGTTAGGA GTTGCAAGCA ATGGACAAGA AAGAGCATT A GGGAAACCAG AATCCCAGCC 1920
TGTCCCTGTG CCTCCACCCC AGCTAGACTT GAGGCTTAGG TCCTGATACT CTGCCCTTAC 1980
ACACCCTCTC TCCATTTCCC AGGAGTTAGT AAAAAACAGG CAAGGGTAAT ATGCTATCAT 2040
CTGTGGTCCT GAGCCTTCTC CAGACTTCCC CTTCACTCCG CCGCCACTGC CACCACCACA 2100
GAGACATCTT CGGAAGCATT AGGTTTATTT CTTGAAC TGA ACTCCTGGGG AGAAGATGAC 2160
TGGTGCCACC ACATGTGCTC TGTCTGGTTA GGGGCCTTGA GGTGGCACAG TCCAGCCATA 2220
GAAGGATCAG CCTTCCTGCC CATAAGATT AATGCCAGG AGTCAGATAG TTGATGCGAC 2280
TGTAAGAGA AAGAGAGGTG AGATGGGGAC TCCAAGGTAC TGTAAGAGA AATCAAGTTT 2340
TTCTTCCCTA GCTCCATCCT CCACTCCAAA GCCTCAGGGA TCTGAAAACA GTCTTCATCC 2400
AGTCTATGTG CCCACAACCT TCAATGGTTC CCTACTTCCT CTTGTATATT TTTGGCTATT 2460
CTCTTCACCT CTAATTAGCT CATACTGTTT CCCACTGACT CCCGTCTTTA ATGGCTTCAT 2520
AGGTTTTTCAG TCACGTATGA ATCTACACTG CATCACACAG TTACAAAACC CAGAGTTGGG 2580
GTGGGGGATT TAGCTCAGTG GCCTAGCAAG CGTAAGGCC C TGGGTTCCGT CCCCAGCCCC 2640
GAAAAA 2646

```

## (2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4339 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW034
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 91..837
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 3101..4162
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 3561..3862
  - (D) OTHER INFORMATION: /label= SAC\_23897

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```

GTCGACGGAA GAGCTGTTGT GACTGCGGTA CTGCGGAGGG CTCTCGGCGG CCCGCGACAA 60
GGCTGAGTCC CGGGAGGTCC ATGCCCCACC ATG CTT TCC TGT GAC ATC TGT GGT 114

```

Met Leu Ser Cys Asp Ile Cys Gly																	
15																	
GAA	ACT	GTA	ACC	TCA	GAA	CCA	GAT	AGG	AAG	GCT	CAC	CTA	ATT	GTT	CAC	162	
Glu	Thr	Val	Thr	Ser	Glu	Pro	Asp	Arg	Lys	Ala	His	Leu	Ile	Val	His		
101520																	
ATG	GAA	AAT	GAA	ATT	ATC	TGT	CCG	TTT	TGC	AAG	CTG	TCG	GGT	ATA	AAT	210	
Met	Glu	Asn	Glu	Ile	Ile	Cys	Pro	Phe	Cys	Lys	Leu	Ser	Gly	Ile	Asn		
25303540																	
TAC	AAT	GAA	ATG	TGT	TTT	CAT	ATT	GAA	ACC	GCT	CAT	TTT	GAA	CAG	ACT	258	
Tyr	Asn	Glu	Met	Cys	Phe	His	Ile	Glu	Thr	Ala	His	Phe	Glu	Gln	Thr		
455055																	
ACG	CCA	GAA	AAA	AGC	TTC	GAG	ACG	CTA	GCC	GCA	GTG	CAA	TAT	GAA	AAT	306	
Thr	Pro	Glu	Lys	Ser	Phe	Glu	Thr	Leu	Ala	Ala	Val	Gln	Tyr	Glu	Asn		
606570																	
TCA	GAC	CTC	GGT	AAT	ACC	AAG	CTG	CAC	AGT	ACA	GTG	GAA	GTT	ACC	TCA	354	
Ser	Asp	Leu	Gly	Asn	Thr	Lys	Leu	His	Ser	Thr	Val	Glu	Val	Thr	Ser		
758085																	
GGC	ATC	CAT	TCA	GCT	TGT	GCA	TCG	AAC	TTT	CCA	AAG	GAG	TCG	TCC	GAA	402	
Gly	Ile	His	Ser	Ala	Cys	Ala	Ser	Asn	Phe	Pro	Lys	Glu	Ser	Ser	Glu		
9095100																	
AGC	CTT	TCT	AAA	GAT	AGG	ACT	TTA	AAA	CAA	GAA	GCT	TTT	TAT	ACA	GAG	450	
Ser	Leu	Ser	Lys	Asp	Arg	Thr	Leu	Lys	Gln	Glu	Ala	Phe	Tyr	Thr	Glu		
105110115120																	
AGC	GTA	GCT	GAA	TCT	AGA	AAG	TAC	CAG	AAA	AGC	AGA	GAA	AAG	CAG	TCC	498	
Ser	Val	Ala	Glu	Ser	Arg	Lys	Tyr	Gln	Lys	Ser	Arg	Glu	Lys	Gln	Ser		
125130135																	
GGA	TTG	TCT	GAG	GCC	CAA	GGA	TCA	ATT	TAT	GAA	ACA	ACA	TAC	AGT	CCC	546	
Gly	Leu	Ser	Glu	Ala	Gln	Gly	Ser	Ile	Tyr	Glu	Thr	Thr	Tyr	Ser	Pro		
140145150																	
CCC	GAA	TGT	CCG	TTC	TGT	GGG	AGA	ATT	GAG	AGG	TAC	AGT	CAA	GAT	ATG	594	
Pro	Glu	Cys	Pro	Phe	Cys	Gly	Arg	Ile	Glu	Arg	Tyr	Ser	Gln	Asp	Met		
155160165																	
GAA	ATT	CAT	GTG	AAA	ACA	AAG	CAT	GCC	AGC	CTT	CTA	GAA	AGT	CCG	TTA	642	
Glu	Ile	His	Val	Lys	Thr	Lys	His	Ala	Ser	Leu	Leu	Glu	Ser	Pro	Leu		
170175180																	
GAA	GAC	TGT	CAT	CAA	CCA	CTC	TAT	GAC	TGT	CCC	ATG	TGT	GGG	CTT	GTC	690	
Glu	Asp	Cys	His	Gln	Pro	Leu	Tyr	Asp	Cys	Pro	Met	Cys	Gly	Leu	Val		
185190195200																	
TGT	ACA	AAT	TAT	CAC	ATT	CTC	CAA	GAA	CAT	GTG	GAC	TTG	CAT	TTA	GAA	738	
Cys	Thr	Asn	Tyr	His	Ile	Leu	Gln	Glu	His	Val	Asp	Leu	His	Leu	Glu		
205210215																	
GAA	AGC	AGC	TTT	CAA	CAA	GTA	CTT	GTG	AGT	GGG	TAT	GTG	CCT	GTG	TGG	786	
Glu	Ser	Ser	Phe	Gln	Gln	Val	Leu	Val	Ser	Gly	Tyr	Val	Pro	Val	Trp		
220225230																	
ACA	CCC	GAC	CTG	CGA	TGT	CAT	CCT	CAG	GAA	GGC	CGT	ACA	ACT	CCT	TTA	834	
Thr	Pro	Asp	Leu	Arg	Cys	His	Pro	Gln	Glu	Gly	Arg	Thr	Thr	Pro	Leu		
235240245																	

AGA TGAGTTTCTT AGTATCTTGG AACACACCAG TTAAGCTCCC AAGTTTGCTT Arg	887
CTATCTCCAC CTCCCTAATA ATGCTGGAGT TACAATCATG GACAACCATA CACACATGTT	947
TCACATGATT CTGGGCACTG AGCTTAGGTC CTTATGTCTG TGAGGCAAGC TCTTCCTGAC	1007
CGAGCTATCT CCTGAGCCTG CTGTCTGTCT CTTTAAATAA TATTAAATAC TTCATATAGA	1067
GCTCTCGAAG GTATAGGGTC TTGTATGAGT CTCCTAGGGC CACTGAAAAA TAAAGTGACC	1127
TCTAAGAGAA GTTGTATTATA TTTTATTAC TCTGTCAGTA TTTGGGAGTA TGTAGACAGC	1187
TCCTGGTAAA GTAGGAGATT TTAGCGTTCG ATGATTTACC ACCAATAACC GTGATCTCTA	1247
TGCATCATGT ACCTGCTTTG CTCATTACCA CAAGAGGAAA ATTAGTAGAA TATTCTTTTT	1307
TTTCCTTTTC CTTTTTTTTTC TTTTTTTCGG AGCTGGGGAC CGAACCAAGG GCCTTGCACT	1367
CGCTAGGCAA GCGCTCTACC ACTGAGCTAA ATCCCAACC CCAAATTAGT AGAATATTGA	1427
TGGAAGATTA TTTCACGTCA GTGGTCTTCT GCTGTCTACC CCAGTTGGCA TGCCACTGTC	1487
GTTGTTTAGG GTTCCTGTTG CTGTAATAAA AACTACCAA CAAGAGTAAC CTGGGCAGGA	1547
AAGGGTTTAT TTTATCCAAA CATTTCCAGT TCACCACTGA GGAATTCAG GGAAGGGACT	1607
CAAGGCAGGA ACCTGTAGGC AGGAGCTGAT GCAAAGGCCA TGAAGTGCTG CTGAGTGGCT	1667
TGTTTTGCCC TGACTTGTTT AGACTCCTTT CTTAGAGCAC CCAGGGATAT CAGCCCAGGG	1727
GTGGCATCAC CTACAAGGAC TGGGCCTTCC AACACCAGTC ACTACAGAAC GCACTACAGA	1787
CTTGCCTGCA CTCGGTCTTT CTTTACTGTT GTATTGTTTT TTGAGATGGA GTCTCACTAT	1847
GTAGCTCTGG CTAAGTGTA ACTCTAGCTT GTGTCAAGTT GACCTAAAC TAGCCAGCAC	1907
AGCTATCTTC CTTTGAGTAA GATTGTAAGA GGATGCTGTG GAAGGGGTAA TTTTTTCAA	1967
GCAGAATGTC TGATAATTCT TACTTTAGGG ATAGGCTTTG TATTGTATTC TATAAACCAA	2027
TGAAACAGTC TCAGGGCATG CTAAAAATCA GTTGTGCAGT ACATGCCCCA GAGTGGAAC	2087
TAGAACCAGT GGAAGGTGT AAAATGGGTT AATCAACAGG TGTGTGCATT GTACTGTCTC	2147
TGAATTGAGG AGTACCTAGA AGACTCCTAA CTGTTTCTTC TGGGTGTATC TCTGAGGATG	2207
TTTTCAGAAT AACTGATGGG TCGGTTGGCA AACAGAATGA AGAAGATGTG CATTGAGCGT	2267
GGGCAAGTAG TGTCAGTGT CGCTAGTACT ATAGGTGGAC CAGGAAAGCC GAAGCAGTGG	2327
CAGAATCACT TTGGCTCCTC CTTCTGGATG GGGCTTTCTT TCCCTGCAGC TGCTCCCTGG	2387
GTGCAGCTGC TCCCTGGGTT ATACTCCAGG ATGTCTGACT TTTGGATGCT GGAACCTGA	2447
TAATCCCTAC CAGAAATTTT GAGCCTTTGA CCTCACACTG GGATATCTGA TTTTGCCAAA	2507
CTGGTTCTGG CACTTTCAAC TTTCATACT GGTTCCTCTG TGTCTCAGGC TATCAAGAAC	2567
TTTTCATTTG TATCATTATG TGATCTATTC CCGTATGATT CGTGATATTT AATAAGGTTT	2627
TCTATATCTT TCTGTCTGTC TGTCTCTGTG TGTGTGTGTA CATAACACA CACGCCTGTG	2687

TGTGGTGT	TTTCTCTTG	CCCTCAGTCC	TGTTACTGGA	GGACCCTGAC	CAGTAAAGAA	2747
TTTAGGAAAC	AAGACTCAAT	ACATTACAAT	ACAATACAAG	CCATTGACTA	AAGAATCTGA	2807
GTTTCAGTGG	GAGAGAGGAA	GCATGGCATT	ATGCTATAAA	ATTATTCAGC	TTGTTCCGGT	2867
AAAAACTGTA	ATGTAATGAA	ACAGTTGGAG	GTTGTGAATA	TATATTTTAT	ATTATTTATT	2927
ACTTTTTTCT	TTCCATTAGT	TGTGTTTCAGA	GATATCAATG	ATAAGGGTCT	TTCAACCGCT	2987
CAECCTTTGT	TTTTCTCCCT	AAAGGTGTTA	GGATGTTTAC	AATTGCTAGT	TAATGAGAAA	3047
TGCTTGTTAG	AGCCTAAGCA	CTGACTCATT	CGGGTTGTCC	CTCTCCAATA	GGC ATG	3103
					Met	
					1	
GAC AGA GTC	CAG TGT TCT	AGT GAC AGA	GAA TTA GCT	CAC CAG CTT	CAG	3151
Asp Arg Val	Gln Cys Ser	Ser Ser Asp	Arg Glu Leu	Ala His Gln	Leu Gln	
	5		10		15	
CAA GAA GAA	GAG AGA AAG	AGG AAA TCT	GAA GAA TCA	AGA CAA GAA	AGG	3199
Gln Glu Glu	Glu Arg Lys	Arg Lys Ser	Glu Glu Ser	Arg Gln Glu	Arg	
	20		25		30	
GAA GAG TTT	CAG AAA TTG	CAG CGG CAG	TAC GGC TTA	GAT AAT TCT	GGA	3247
Glu Glu Phe	Gln Lys Leu	Gln Arg Gln	Tyr Gly Leu	Asp Asn Ser	Gly	
	35		40		45	
GGA TAC AAA	CAA CAG CAG	CTG CGA CAC	ATG GAG CTA	GAA GTA ACG	AGG	3295
Gly Tyr Lys	Gln Gln Gln	Leu Arg His	Met Glu Leu	Glu Val Thr	Arg	
	50		55		60	65
GGA CGA ATG	CAT CCA TCT	GAG TTC CAC	AGC AGA AAA	GCT GAC ATG	TTG	3343
Gly Arg Met	His Pro Ser	Glu Phe His	Ser Arg Lys	Ala Asp Met	Leu	
	70		75		80	
GAA TCA ATC	GCT GTT GGT	ATT GAT GAT	GGA AAA ACA	AAA ACT TCT	GGA	3391
Glu Ser Ile	Ala Val Gly	Ile Asp Asp	Gly Lys Thr	Lys Thr Ser	Gly	
	85		90		95	
ATT ATT GAA	GCC CTC CAC	AGG TAT TAT	CAG AAC ATT	GCC ACA GAT	GTA	3439
Ile Ile Glu	Ala Leu His	Arg Tyr Tyr	Gln Asn Ile	Ala Thr Asp	Val	
	100		105		110	
AGG TGT GTG	TGG CTG TCT	ACA GTG GTG	GAT CAC TTT	CAT TCA TCT	TTT	3487
Arg Cys Val	Trp Leu Ser	Thr Val Val	Asp His Phe	His Ser Ser	Phe	
	115		120		125	
GGG GAC AAA	GGT TGG GGT	TGC GGT TAT	AGA AAT TTC	CAG ATG CTG	CTT	3535
Gly Asp Lys	Gly Trp Gly	Cys Gly Tyr	Arg Asn Phe	Gln Met Leu	Leu	
	130		135		140	145
TCA TCA TTA	CTG CAG AAT	GAA GTG TAC	AGT GAC TGC	TTG AAA GGT	ATG	3583
Ser Ser Leu	Leu Gln Asn	Glu Val Tyr	Ser Asp Cys	Leu Lys Gly	Met	
	150		155		160	
TCA GTT CCT	TGT ATT CCA	AAA ATT CAG	TCC ATG ATT	GAA GAT GCG	TGG	3631
Ser Val Pro	Cys Ile Pro	Lys Ile Gln	Ser Met Ile	Glu Asp Ala	Trp	
	165		170		175	
AAT GAA GGT	TTT GAT CCT	CAG GGG GCC	TCT CAA CTT	AAT AAC AAG	TTG	3679
Asn Glu Gly	Phe Asp Pro	Gln Gly Ala	Ser Gln Leu	Asn Asn Lys	Leu	

180	185	190	
CAG GGG ACC AAG GCC TGG ATT GGA GCA TGT GAG ATC TAT ACA CTT CTG Gln Gly Thr Lys Ala Trp Ile Gly Ala Cys Glu Ile Tyr Thr Leu Leu 195 200 205			3727
ACC TCA CTG AAA GTC AAG TGC CGC ATT ATT GAT TTT CAC AAG TCA ACT Thr Ser Leu Lys Val Lys Cys Arg Ile Ile Asp Phe His Lys Ser Thr 210 215 220 225			3775
GGT CCT TCA GGC ACA CAC CCT CGC TTA TTC GAG TGG ATA CTG AAC TAT Gly Pro Ser Gly Thr His Pro Arg Leu Phe Glu Trp Ile Leu Asn Tyr 230 235 240			3823
TAT TCC TCA GAG ACG GAA GGG GCG CCA AAG GTC GTG TGT ACA TCC AAA Tyr Ser Ser Glu Thr Glu Gly Ala Pro Lys Val Val Cys Thr Ser Lys 245 250 255			3871
CCT CCT GTC TAT CTC CAG CAT CAG GGT CAC AGT CGA ACA GTT GTT GGG Pro Pro Val Tyr Leu Gln His Gln Gly His Ser Arg Thr Val Val Gly 260 265 270			3919
ATT GAA GAG AGA AAA AAC CGA ACA TTA TGT TTG CTA ATA TTT GAT CCT Ile Glu Glu Arg Lys Asn Arg Thr Leu Cys Leu Leu Ile Phe Asp Pro 275 280 285			3967
GGA TGT CCT TCT CGA GAA ATG CAG AAG CTG TTA AAG CAA GAC ATG GAG Gly Cys Pro Ser Arg Glu Met Gln Lys Leu Leu Lys Gln Asp Met Glu 290 295 300 305			4015
GCT GGT AGT CTC AGG CAG CTC CGG AAA TGT GTG GGA AAT CTG AAG CAT Ala Gly Ser Leu Arg Gln Leu Arg Lys Cys Val Gly Asn Leu Lys His 310 315 320			4063
AAG CAG TAC CAG ATA GTA GCA GTG GAG GGT ATC CTG TCC CCA GAG GAG Lys Gln Tyr Gln Ile Val Ala Val Glu Gly Ile Leu Ser Pro Glu Glu 325 330 335			4111
AGA GCT GCC AGG AAA CAA GCT TCT CAA GTG TTT ACA GCG GAG AAG ATT Arg Ala Ala Arg Lys Gln Ala Ser Gln Val Phe Thr Ala Glu Lys Ile 340 345 350			4159
CCT TGACCAGAT ATTTTGTAAT TATCCTTTTT GTTCCGATA TTGAACTCTG Pro			4212
ATACAGTTGA AGAATTTGAC TTCACTGAAG TCGGTGATAA ATTATTTAAG TTATAAATGC			4272
CTGTTGTACT CCTTAGCATT TCGTTGTCTG GTAATCAGTT GAATAAATTC CTGTTTACA			4332
TTAAAAA			4339

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 249 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Leu Ser Cys Asp Ile Cys Gly Glu Thr Val Thr Ser Glu Pro Asp  
 1 5 10 15  
 Arg Lys Ala His Leu Ile Val His Met Glu Asn Glu Ile Ile Cys Pro  
 20 25 30  
 Phe Cys Lys Leu Ser Gly Ile Asn Tyr Asn Glu Met Cys Phe His Ile  
 35 40 45  
 Glu Thr Ala His Phe Glu Gln Thr Thr Pro Glu Lys Ser Phe Glu Thr  
 50 55 60  
 Leu Ala Ala Val Gln Tyr Glu Asn Ser Asp Leu Gly Asn Thr Lys Leu  
 65 70 75 80  
 His Ser Thr Val Glu Val Thr Ser Gly Ile His Ser Ala Cys Ala Ser  
 85 90 95  
 Asn Phe Pro Lys Glu Ser Ser Glu Ser Leu Ser Lys Asp Arg Thr Leu  
 100 105 110  
 Lys Gln Glu Ala Phe Tyr Thr Glu Ser Val Ala Glu Ser Arg Lys Tyr  
 115 120 125  
 Gln Lys Ser Arg Glu Lys Gln Ser Gly Leu Ser Glu Ala Gln Gly Ser  
 130 135 140  
 Ile Tyr Glu Thr Thr Tyr Ser Pro Pro Glu Cys Pro Phe Cys Gly Arg  
 145 150 155 160  
 Ile Glu Arg Tyr Ser Gln Asp Met Glu Ile His Val Lys Thr Lys His  
 165 170 175  
 Ala Ser Leu Leu Glu Ser Pro Leu Glu Asp Cys His Gln Pro Leu Tyr  
 180 185 190  
 Asp Cys Pro Met Cys Gly Leu Val Cys Thr Asn Tyr His Ile Leu Gln  
 195 200 205  
 Glu His Val Asp Leu His Leu Glu Glu Ser Ser Phe Gln Gln Val Leu  
 210 215 220  
 Val Ser Gly Tyr Val Pro Val Trp Thr Pro Asp Leu Arg Cys His Pro  
 225 230 235 240  
 Gln Glu Gly Arg Thr Thr Pro Leu Arg  
 245

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Met Asp Arg Val Gln Cys Ser Ser Asp Arg Glu Leu Ala His Gln Leu  
 1 5 10 15



WO 98/53071

Gln Gln Glu Glu Glu Arg Lys Arg Lys Ser Glu Glu Ser Arg Gln Glu  
 20 25 30  
 Arg Glu Glu Phe Gln Lys Leu Gln Arg Gln Tyr Gly Leu Asp Asn Ser  
 35 40 45  
 Gly Gly Tyr Lys Gln Gln Gln Leu Arg His Met Glu Leu Glu Val Thr  
 50 55 60  
 Arg Gly Arg Met His Pro Ser Glu Phe His Ser Arg Lys Ala Asp Met  
 65 70 75 80  
 Leu Glu Ser Ile Ala Val Gly Ile Asp Asp Gly Lys Thr Lys Thr Ser  
 85 90 95  
 Gly Ile Ile Glu Ala Leu His Arg Tyr Tyr Gln Asn Ile Ala Thr Asp  
 100 105 110  
 Val Arg Cys Val Trp Leu Ser Thr Val Val Asp His Phe His Ser Ser  
 115 120 125  
 Phe Gly Asp Lys Gly Trp Gly Cys Gly Tyr Arg Asn Phe Gln Met Leu  
 130 135 140  
 Leu Ser Ser Leu Leu Gln Asn Glu Val Tyr Ser Asp Cys Leu Lys Gly  
 145 150 155 160  
 Met Ser Val Pro Cys Ile Pro Lys Ile Gln Ser Met Ile Glu Asp Ala  
 165 170 175  
 Trp Asn Glu Gly Phe Asp Pro Gln Gly Ala Ser Gln Leu Asn Asn Lys  
 180 185 190  
 Leu Gln Gly Thr Lys Ala Trp Ile Gly Ala Cys Glu Ile Tyr Thr Leu  
 195 200 205  
 Leu Thr Ser Leu Lys Val Lys Cys Arg Ile Ile Asp Phe His Lys Ser  
 210 215 220  
 Thr Gly Pro Ser Gly Thr His Pro Arg Leu Phe Glu Trp Ile Leu Asn  
 225 230 235 240  
 Tyr Tyr Ser Ser Glu Thr Glu Gly Ala Pro Lys Val Val Cys Thr Ser  
 245 250 255  
 Lys Pro Pro Val Tyr Leu Gln His Gln Gly His Ser Arg Thr Val Val  
 260 265 270  
 Gly Ile Glu Glu Arg Lys Asn Arg Thr Leu Cys Leu Leu Ile Phe Asp  
 275 280 285  
 Pro Gly Cys Pro Ser Arg Glu Met Gln Lys Leu Leu Lys Gln Asp Met  
 290 295 300  
 Glu Ala Gly Ser Leu Arg Gln Leu Arg Lys Cys Val Gly Asn Leu Lys  
 305 310 315 320  
 His Lys Gln Tyr Gln Ile Val Ala Val Glu Gly Ile Leu Ser Pro Glu  
 325 330 335  
 Glu Arg Ala Ala Arg Lys Gln Ala Ser Gln Val Phe Thr Ala Glu Lys  
 340 345 350

Ile Pro

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2415 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

(B) CLONE: HW035

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1738..1910)
- (D) OTHER INFORMATION: /label= SAC\_24468

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GTCGACGGGC CCTATTCTGC CTCCTTCAGG ATCTGCGTCT TCCAGGGATC CGCTGTTCTGA	60
CTTGAGAGTG CTTCTCTGCA GGTCCCAAGC CTGCCTCCTC ACCCCAGCTT TCCAGTGGAG	120
GCTGCCCCCT TCCAGGCTTG GCTCAGGAAG CTCCTCATAG GCGGATGCCC ACCCATAAGA	180
GGCTACATCC CTTTGTGGTC ACATGGCCTG TAGCCTCTTG CTGACATTGC TGGTGTCTTC	240
TCTGGGGCCC TGGATCCCTG ACAAGCACAG AAATGTGATC TGGTAAAAGG AAAGACTGTG	300
GCCTGATAGG CCTTTACCTC GTAGGCAGCA GGATTTTCAC TCAGTTTGAC ACAGCATGCA	360
GTGCTCTGGG AGCAGGACCT CTGCCACCTT GGCATTTTGT GTAGGCTGGT TTGGCTCAAG	420
GTGCTCAGAC ACTCCCATTG ATGCCTGAGT CGGATATGGG TGGAAATCTG GACTGTTGAT	480
TATAAGTTAC CAGGTGACCT CATTCCTCC CTTTGTCTCT TAACTGGTAT CCAGTGAGTT	540
GCTGATGCCG ACTGGCAAGC AGAGCTGGCC CATGCCTAGA GGAGGAGAGG GACAGGCTTT	600
CTGAGTGCTT TGATTCTGGC TGGGCTGTAG TGTATAGCTG AGAATGGCCT CCTGCATGGG	660
TCCAGCCTTT TGGGAGGATG AGGAAGAGGA GGCTGAGGCT AGATGTGTGT TAAACAGGAG	720
TAGCAGATCT TTGCTTCTGA ACCGGATAGA GCCGTGCTCA AGGC'TTTCAG GTTTGCAGTA	780
TCTTCTTATT GAGTCTGGGG AACGTTCCCT ATTGAGCCAG CCCTGCAGCT CCACTCCCTG	840
CCCCATTCCC CTGATTCTTT GACAGTAAAT CCTGTGTCCA GTTGTGACTG TTGGCTGTTT	900
TCCTGATCAT TGTTAATTCC GTTTTCCCTAC TGTGAGGGTG TTGTGTTTGA CTTTCTGACA	960
CAGCGAATGG AGAGCTTGGG CTGAGGCCAG GTCTCCCAA CTTCTCCGCT AAGTAAACCG	1020
TAATGTGTGA CTTCCCTCCT ATCTTGACTG CCCGACTCCT TCCTTCCACA TGGACACTCA	1080

GATTTCAAAG AAAACCCGCA GGAGGTGGGT CAGTTTCCTC TGTAAGATAT GGTCCCAGAG	1140
AGTCCTCCCC TCTGTTGTGT CTAGCTGCTA GTCTGGGTGT TGTACCATCA GTGCAGGGAG	1200
TCCACGGCAA CACTGAGTTA TACAGGGACC ACTACCATCG CCAGGCTTCC TAGCAGGCTG	1260
TTCTCAGTTC ATCTTTCAGG TCCGTGTGTA CTCTGTGCTA GCAGCAGCCT TGGTGTTCAA	1320
GGCCAGGATT TGCAGTTTGT ATAAGCTACA GGGCACTTGG AGATGGTGAG GCCACTTCCC	1380
TGCTACCTCC CTCTCTGGTG TGGCCACTTG GCCAAAGCTC CCATCCACAG CTGGGGTGTC	1440
TGAGCCACAG CCTGCTACAC TGTTCGCTCT CTGGGATCTT CTGAGTTACT CTTTTAGGCT	1500
GTGTTTGGT GAAAGGAACC AACACATTAA CGATCTCCCC CCCCCCCCCC AAAGCCACTG	1560
AGTAATTCTT GGCATGTTTT CTAAGTGCTG TTGGCTGCCT GGTTCGGGAG CAGTTGTGGA	1620
GAGGACAGAC TTAGTACCAC ACAGAGGCAT GGGTGTGCGA GCTGCTACA GCTTCCTTGT	1680
TTCTGGCATG TGTCCTGGGT GCGTGCTATA CCTCCCGGGC CAGCTGCACA CCTATGTACT	1740
CTGCCCCGA GAGGTAGAGT GTTCAGGAAG TGTCCCTGAT AGATCCCTG TTCTCTGTGG	1800
CCCCAGTTGG GCTTTTGGCT CCTCCAGCAA CTTTCATGCC CAACCCTACC TTCTCTCCCA	1860
TTTCTCTCTG TCTGCTAGTC CCTGAAGCCT TTAACCAAAT GGAATGGGT ACAGAAAGCC	1920
TTCTCCTGGC AACTTAGGGC AACAGGACAA GGGCTATTGT GTGTCCAACCT CTGTTGGTCT	1980
CCTGGTGCTG AGAGGTGTGT CCAAGCAGAG TTGATCAGCC CCTGCCTGCC CTGCAGGGCT	2040
GAAGCCAGGG GGAGGTAGT AGAGAGTACC CTAAATGGGG TGAGGGTCCT TAAACATCCA	2100
AGTGAGGAGG GGCTGTTGTG TGGCTGCAAG GTGAGACCCT GTCTCCAGCA CCTGGGTAGG	2160
GCCATCGGGG ATCTGGCTGC AAGGTGAGGT CTGGACTTGG CATCAGTGCA GAGCGATCAA	2220
GTCACCTGGC CTCAGCAGGC AGCAGCCTTG GTGCCCCACC GTGCCCTGCG CAGTATTTAT	2280
TGCTAAATTA TTGTCCAGGA GGGGCGGCAC TGGGCCTGGC CCCCCGGGTA TTTATTGCTG	2340
TACATAGTGT ATGTTTGTGA TATATAAGGT TTTCTTTATT TTGTATATGA TCAATAAACG	2400
CCTTTTAAAG AAAAA	2415

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3531 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW036

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION: complement (2541..2575)  
 (D) OTHER INFORMATION: /label= SAC\_24406

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTCGACTGAA AACCACCTCAG AATTTTCTCT GGCCTTAAC AGAAAAGAAT CAAAGTAGGA	60
AGCACAGTAG CTGCCACATC TGTTC AACCT CTGGGCCTCC CAAGTACACC AGAGAAAACA	120
GATCAATGCT GTTCTGAAT GATTCTGCTG GTTAAAGTCA TTTATAGAAG ATTACAGATA	180
CTGAACAACCT ATGGGTAAAG AACTTACTGT GTTTTGTGAA TTGCCCAATG GAGAGAGTGT	240
GTTCAAGGAT GACTCTTGAT CAGATTACAC AGTATGTGCA TCCAGGAGGC CACATGAAAT	300
CAGCATAGTT AGTAACATAG AATCTGTTGT CAGCCTCCTC CAACTCGTGT CTGTTTCTTG	360
GTTTCAATTT TAGGCCAGCC CTCTCCTCAT AAAGGGGTGA TAGTAGCCAC AGTCCACAGC	420
CTCAAGGTCT GGTCTCCAAA AAGAAGAGTG TGAGCTGGGC CACAATCATA CCCTCCTGAA	480
GAAGTCTGGT CTTTTCTGGT CATATGCTCC TGCCCCACCC TTGGTGATCC TGGCTAGATC	540
ATGGTCTCCC GTACTGTGCC GACACCATTC ACAGTAACAT AGGATGAGTT ATGGGCCTCA	600
TCCCTAAACA CTGGAACCTG CAGCAAATGA TGTAGCAGAT ATTCACTCTA GATAACAAAT	660
GTCAGTGACA TGAAGTGCT AGGAAAGTTT AGAACCTGA TCATTTTATT TTTACCACTT	720
AACACAGTAG AGAGCCATCA GGACTAAATC TTTGGTGTGA TTCTTGCCAT TGAAGTCATT	780
ATAGCTTCCA TATTGTCCG TAGTTCACTG GAAAGTGCAC TTGCTGTAA AAGTTTTGAA	840
GCTGTGAGGA TTCCCTAGGG TGTCAC TGAA GTCTGCTCTG TATACTGCCC CCACTGGGAG	900
TGCACGGTAC TGCACCCTTA AATCCTGCCG CCTGTGTCCC AACATGGTG TCTGCGTGCA	960
GTGTGGAGTA GTGCGCCTTC TTTTTTTTTT TTTTTTTTTT CCTATTCTTT TTTTGGAGCT	1020
GGGAACCGAA CCCAGGGCCT TGTGCTTGCT AGGCAAGTGC TCTACCACTG AGCCAAATCC	1080
CCAACCCCTG TAGTGC GCCT TCTATACTAG AAAGCTTGAC CACTGAGCCA CACCTCCCAC	1140
TAGTGCTTCA ATGTCAACCG AGAGTAAAT GTGTTTGAT GAAATGCCTC CATTTGACTA	1200
GATAGAGCTT TATTTGGAGA AAGTCACATA TAACATAATT GAACTTTGAA TTATACAATC	1260
CCGTGGATTT TAGAGTGCTC CTGGAGCAGG TGGCAGTCAC CACTATCTAC TTCCAGAACA	1320
GTCTCATCCT TTCCAGAAAC CCACACTCTG TCTTTCCTCT ATTCCAGATC TGTTAGACGA	1380
GTGGAATTAC ATAGTCCGGT CTTTCTGAG TTCTGTACT AAGTTTAAA GGTTATTCT	1440
CAGGTAGCAT CAGTCCGTAA TGTATTACTG CTGAATAGTG TTCCGTGTAT ACAGACACCG	1500
TGTGTGCTT CTCCAGCGA GCAGAGGAAC TCTGAGCTGT TTCTACTTTG GGGCTTTTGA	1560
CTAATGCTAT GAACATCTGT GAAAAAGTTC GAAATGTTTG ATTTAGTACA GACCCTAGTG	1620
GGGAGCTCCG GGGTCATATT ATGACAGCCT CAATTGTACT TCCTACAGTG GTTTTACCAC	1680

CATTTCTGTC	TCTCGTGGGA	TCTAGGCTCC	AGCATCCCTC	ACAACTTTCT	GCCTGAGATG	1740
AAGAGGCATC	TGATTGGGAT	CTTGGTTTGC	ATTTCCCTAA	TGTCTAATAA	TCTGAGCTTT	1800
TTTTCATGTG	TTCATTGGCT	TTCTATGCTG	CTTTGCAGAA	TGTTTATTTT	AGGCTACAGT	1860
CTGCCTTTCA	GCTGGGTTAT	CTTTCTGTTT	TTCTGTAGGA	TTTTTTATTT	ACGGTCAACT	1920
CATCTCTTAA	AGATTAATTG	GCATTTTTTT	TTTTCTCAAC	TTGTGGTTTG	CCCTTCTCAA	1980
TGTCTTGATG	CTCTCTGAAG	CACAGTTTTT	ATTTTATAT	ATTGCCTTTC	TGCGTCTTAG	2040
CCCAGGCTGG	CCAGGAATCC	CCAAAGCCTT	GTACACAGAG	TGAGCATTTT	CACCGCCTCA	2100
CTTCCCAGAA	GGCACACGGC	GCTGCTCCTG	CCTAGAGCAG	TTTTCATATT	GGAAAATAAC	2160
AATGTGTTTT	TTCTTTCAGT	GTATCCTAAG	AAACCATTCC	TAACCTGGAC	TTACAGAAGC	2220
TTATGTCTTC	ACTAAAGAAT	TTTTAAATTT	TACCTCTTAC	CTGTAGGCCT	GTGCTCTCAA	2280
AATGTTACAC	ACCTTTATTC	ATGGCCATAT	TTAAATTAAT	GTAAGAAATG	TGAGTATGTA	2340
GATATGCCTG	TATCGTGATG	AGCACATTCA	CAGTGCAGTG	CAGTTGGCCT	GCAGCCCAGT	2400
TCTGCATCCA	CCTCAGAAAA	TGCCCAGTGA	GCCCATCTAA	TGACTCCCAG	CCCCTGCCAG	2460
GTCTCCCTC	CCTGCCCTGT	CTTGGCATA	TCTCTCTCC	ATGGCAATCA	CCATTCTCTG	2520
TCCTGTTTCT	GAATTTGAGT	ACCCTAAGTG	CCTGGTATAC	ATAGATTCTC	AGTGCCTGTG	2580
TGTGTGTGTG	TGTGTGTGTG	TGTGCGTGCG	TGTTTATGTA	GCAGAATGCC	GTGAAGGCTT	2640
GTATCATGTA	TGAGTTTCCT	TTGTAAGGAC	AAATAGTACT	TTAGACTTTA	GTGTGTGTG	2700
TATGTATGCT	GTATGTTTCT	TGTTACCTAC	TCATCCAGCC	GTGGGTACCT	CAGTTGTTCC	2760
CACCTTGAG	CTCTGACTA	TGCTGTGAAC	ATGTATATAC	AAATATCCCT	CGGATCCTGC	2820
TTAGCATACA	CGCAGCAGAT	TGTATGTCTT	AGGGTAACTC	TGCTGTTATT	AGTAGGCACT	2880
GCCGTAGTCT	CCATAGTGAC	CACACTATTG	TATGTCTGG	CAGCAGTGCA	CAGAGATTAG	2940
AAGTGCCCAT	GTCTTGTGA	GTCTTACTCT	TTTTTTAGTG	GCTGTTCTAG	TAGGTGTTAA	3000
GATCTGCTAG	TGCTTCTGCT	TCCCATCTCT	CTGACTACTG	ATGTCACATG	TTTTATATGC	3060
TTGGTGGCCT	TCTTTATATA	TCGTTTGAAG	AAATCTCTTT	AGGTCCACTG	CCCATTAAATA	3120
GTTGGGTCGT	TGGCTCTGGT	TGAATTTTTG	GTAGTCACAT	CCACAAAAGC	ATTAACCCAC	3180
CCAAGGGAGC	CTACTACAGC	TTGCCCTAGG	AACTTGATT	TTAGCTCTTA	GCTTTGCTCT	3240
CTGGTATATT	TTGAGTTCAT	TTGTGTGTGA	GATGTAAGGT	CTGACTTTCT	ACTCTACCCA	3300
TGTGGATACC	TGGCTTTGCA	TGTTGCCCTT	GGTGGGAAAA	CATCCCGCTG	TGCATGGTGG	3360
GGCTGTTTCT	GAGCCTGGTC	TCTTCCATGA	CCTCTGTCTG	CCTTCATGCC	AACGCCACTG	3420
CTTTGATTAC	TGGAGAATTA	TAATAGAAAT	TTGAAATCAG	AAAATGTTTG	TTCTGTAATA	3480
AAAATTTGAA	ATCAGAAAGT	GTTCTATAAT	AAAAATTTGA	AATCAGAAAA	A	3531

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2127 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW037

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 280..1422

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1370..1796
- (D) OTHER INFORMATION: /label= SAC\_24354

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GTCGACCTGC GGCTGGCCCG CGGGGCAGGC CCCGGACCCG TGCGGTTTCCT GGGCATGGTG	60
AGCAAGGGGC TGCTGCGCCT TGTCTCTTCA GTCAACCGCA GGAAGATGAA GCTTCTGCTG	120
GGCATCGCGC TGTTGCGCTA CGCCGCCTAA CTTGCTCATG ATGGAGTGCC AAGTGCCTCA	180
GAGAAACAGG CCACCAGAAG AACAGTGCAT TCTCAGTACA AGGGCAGCGT TCAGCATTTG	240
GATATTGGAG TTGAAGCTGT TTGGGGCAAC TTTGTTAAC ATG AGG TCT ATC CAG	294
Met Arg Ser Ile Gln	
1 5	
GAA AAT GGT GAA CTA AAG ATT GAA AGC AAG ATT GAA GAG ATT ATT GAA	342
Glu Asn Gly Glu Leu Lys Ile Glu Ser Lys Ile Glu Glu Ile Ile Glu	
10 15 20	
CCA TTA AGA GAG AAA ATC AGA GAT TTG GAA AAA AGT TTC ACC CAG AAA	390
Pro Leu Arg Glu Lys Ile Arg Asp Leu Glu Lys Ser Phe Thr Gln Lys	
25 30 35	
TAC CCA CCA GTA AAG TTT TTG TCA GAA AAG GAC CGG AAG AGG ATT TTG	438
Tyr Pro Pro Val Lys Phe Leu Ser Glu Lys Asp Arg Lys Arg Ile Leu	
40 45 50	
ATC ACT GGA GGT GCG GGC TTT GTG GGC TCC CAT CTA ACT GAC AAA CTC	486
Ile Thr Gly Gly Ala Gly Phe Val Gly Ser His Leu Thr Asp Lys Leu	
55 60 65	
ATG ATG GAT GGC CAT GAG GTG ACC GTG GTG GAC AAC TTC TTC ACA GGC	534
Met Met Asp Gly His Glu Val Thr Val Val Asp Asn Phe Phe Thr Gly	
70 75 80 85	
AGG AAG AGA AAT GTG GAA CAC TGG ATT GGC CAT GAG AAC TTC GAG CTG	582
Arg Lys Arg Asn Val Glu His Trp Ile Gly His Glu Asn Phe Glu Leu	
90 95 100	

ATT AAC CAT GAT GTG GTA GAG CCA CTC TAC ATC GAA GTT GAC CAG ATC Ile Asn His Asp Val Val Glu Pro Leu Tyr Ile Glu Val Asp Gln Ile 105 110 115	630
TAC CAT CTG GCT TCT CCA GCC TCC CCT CCA AAC TAC ATG TAC AAC CCC Tyr His Leu Ala Ser Pro Ala Ser Pro Pro Asn Tyr Met Tyr Asn Pro 120 125 130	678
ATC AAG ACC CTG AAG ACA AAT ACA ATT GGA ACA CTA AAC ATG TTG GGA Ile Lys Thr Leu Lys Thr Asn Thr Ile Gly Thr Leu Asn Met Leu Gly 135 140 145	726
CTG GCA AAG CGT GTG GGT GCC CGT CTG CTC CTA GCC TCC ACA TCT GAG Leu Ala Lys Arg Val Gly Ala Arg Leu Leu Leu Ala Ser Thr Ser Glu 150 155 160 165	774
GTA TAT GGA GAT CCT GAG GTC CAC CCT CAA AGT GAG GAC TAC TGG GGC Val Tyr Gly Asp Pro Glu Val His Pro Gln Ser Glu Asp Tyr Trp Gly 170 175 180	822
CAC GTG AAT CCC ATA GGA CCC CGA GCC TGC TAT GAC GAG GGC AAA CGT His Val Asn Pro Ile Gly Pro Arg Ala Cys Tyr Asp Glu Gly Lys Arg 185 190 195	870
GTT GCG GAA ACC ATG TGC TAT GCC TAC ATG AAG CAG GAA GGT GTG GAG Val Ala Glu Thr Met Cys Tyr Ala Tyr Met Lys Gln Glu Gly Val Glu 200 205 210	918
GTG CGG GTG GCC AGG ATC TTC AAC ACC TTT GGG CCA CGA ATG CAC ATG Val Arg Val Ala Arg Ile Phe Asn Thr Phe Gly Pro Arg Met His Met 215 220 225	966
AAT GAT GGG AGG GTG GTC AGC AAC TTC ATC TTA CAA GCG CTA CAA GGG Asn Asp Gly Arg Val Val Ser Asn Phe Ile Leu Gln Ala Leu Gln Gly 230 235 240 245	1014
GAG CCG CTC ACA GTA TAT GGA TCT GGG TCT CAG ACA CGG GCA TTC CAG Glu Pro Leu Thr Val Tyr Gly Ser Gly Ser Gln Thr Arg Ala Phe Gln 250 255 260	1062
TAT GTT AGC GAT CTA GTG AAT GGC CTG GTA GCA CTG ATG AAC AGC AAT Tyr Val Ser Asp Leu Val Asn Gly Leu Val Ala Leu Met Asn Ser Asn 265 270 275	1110
GTC AGC AGC CCT GTC AAC CTG GGA AAT CCA GAA GAA CAC ACA ATC CTG Val Ser Ser Pro Val Asn Leu Gly Asn Pro Glu Glu His Thr Ile Leu 280 285 290	1158
GAA TTT GCT CAG TTA ATT AAA AAC CTT GTT GGT AGT GGA AGT GAA ATT Glu Phe Ala Gln Leu Ile Lys Asn Leu Val Gly Ser Gly Ser Glu Ile 295 300 305	1206
CAG TTT CTC TCT GAA GCT CAG GAT GAT CCA CAG AAA AGA AAA CCA GAC Gln Phe Leu Ser Glu Ala Gln Asp Asp Pro Gln Lys Arg Lys Pro Asp 310 315 320 325	1254
ATC AAA AAA GCA AAA CTG ATG CTG GGG TGG GAG CCT GTG GTT CCA TTG Ile Lys Lys Ala Lys Leu Met Leu Gly Trp Glu Pro Val Val Pro Leu 330 335 340	1302
GAG GAA GGA TTG AAC AAA GCC ATC CAC TAT TTC CGG AAG GAA CTA GAG Glu Glu Gly Leu Asn Lys Ala Ile His Tyr Phe Arg Lys Glu Leu Glu 1350	

345	350	355	
TAC CAG GCT AAT AAC CAG TAC ATC CCT AAA CCC AAG CCT GCC AGA GTG			1398
Tyr Gln Ala Asn Asn Gln Tyr Ile Pro Lys Pro Lys Pro Ala Arg Val			
360	365	370	
AAG AAG GGC CGG ACG CGC CAC AGC TGAGTTAGCC TTGGGATGCG AGACTCTATT			1452
Lys Lys Gly Arg Thr Arg His Ser			
375	380		
TTACTTTTATA AGGTGGACTT TTGTGGGATT TTTTTTTTTT TTTTAAGACT TAAACAGGTG			1512
TCATGAAGAA CAAACTGGAA TTTTATTCTG AAGCTTGCTT TAAAGACACT GATGTGCCTA			1572
AAAGCTCCCT TGAACCTCTGC AGACTTTGCC TTGCACTTTT TAACTCTGTC TTTTTTATGC			1632
ACAACAGCCT AGATGCATTC TCTGCTATTT TCAGGTTTTT TTATCTTGCT GTTAGAGTGT			1692
ATGCTGTAAC TGTCACGTAC AGTTTTATTT ACTGGTTTCT TTGTGAAGCT GAAAAGAAAC			1752
ATTAAATGGG GTGGAAAATG CCAATTTTAT TTATAAAAGT GAGTACTTTA TAAATGAGAT			1812
GTTACACTAT GCATAAAGAA TACAAACCTA GAGGTATGGC CAGCTGGGCA GTGCACCAGT			1872
GTTTATTTGG GGACAGATGA AAGAACTCCA TTGGAAAGCT TTGCATTCTT TTTTAAATTC			1932
TGAATTTTCT GAAGGTCTAG TTTTCAGTTA CAGACTTGAC TTTGAAACAT CCCTGTTGGT			1992
TCTTGATCAA AGATATTTGA AATCACTACT GTGTTGTGCT GCATATTGGG GTGGGGGTGG			2052
GGACAATGTT AACATATTCT TGGTTAACCA TGGTTAAATA TGCTATTTTA ATAAAATATT			2112
GAAACTCGTC AAAAA			2127

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 381 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Arg Ser Ile Gln Glu Asn Gly Glu Leu Lys Ile Glu Ser Lys Ile			
1	5	10	15
Glu Glu Ile Ile Glu Pro Leu Arg Glu Lys Ile Arg Asp Leu Glu Lys			
20	25	30	
Ser Phe Thr Gln Lys Tyr Pro Pro Val Lys Phe Leu Ser Glu Lys Asp			
35	40	45	
Arg Lys Arg Ile Leu Ile Thr Gly Gly Ala Gly Phe Val Gly Ser His			
50	55	60	
Leu Thr Asp Lys Leu Met Met Asp Gly His Glu Val Thr Val Val Asp			
65	70	75	80
Asn Phe Phe Thr Gly Arg Lys Arg Asn Val Glu His Trp Ile Gly His			
85	90	95	



Glu Asn Phe Glu Leu Ile Asn His Asp Val Val Glu Pro Leu Tyr Ile  
 100 105 110  
 Glu Val Asp Gln Ile Tyr His Leu Ala Ser Pro Ala Ser Pro Pro Asn  
 115 120 125  
 Tyr Met Tyr Asn Pro Ile Lys Thr Leu Lys Thr Asn Thr Ile Gly Thr  
 130 135 140  
 Leu Asn Met Leu Gly Leu Ala Lys Arg Val Gly Ala Arg Leu Leu Leu  
 145 150 155 160  
 Ala Ser Thr Ser Glu Val Tyr Gly Asp Pro Glu Val His Pro Gln Ser  
 165 170 175  
 Glu Asp Tyr Trp Gly His Val Asn Pro Ile Gly Pro Arg Ala Cys Tyr  
 180 185 190  
 Asp Glu Gly Lys Arg Val Ala Glu Thr Met Cys Tyr Ala Tyr Met Lys  
 195 200 205  
 Gln Glu Gly Val Glu Val Arg Val Ala Arg Ile Phe Asn Thr Phe Gly  
 210 215 220  
 Pro Arg Met His Met Asn Asp Gly Arg Val Val Ser Asn Phe Ile Leu  
 225 230 235 240  
 Gln Ala Leu Gln Gly Glu Pro Leu Thr Val Tyr Gly Ser Gly Ser Gln  
 245 250 255  
 Thr Arg Ala Phe Gln Tyr Val Ser Asp Leu Val Asn Gly Leu Val Ala  
 260 265 270  
 Leu Met Asn Ser Asn Val Ser Ser Pro Val Asn Leu Gly Asn Pro Glu  
 275 280 285  
 Glu His Thr Ile Leu Glu Phe Ala Gln Leu Ile Lys Asn Leu Val Gly  
 290 295 300  
 Ser Gly Ser Glu Ile Gln Phe Leu Ser Glu Ala Gln Asp Asp Pro Gln  
 305 310 315 320  
 Lys Arg Lys Pro Asp Ile Lys Lys Ala Lys Leu Met Leu Gly Trp Glu  
 325 330 335  
 Pro Val Val Pro Leu Glu Glu Gly Leu Asn Lys Ala Ile His Tyr Phe  
 340 345 350  
 Arg Lys Glu Leu Glu Tyr Gln Ala Asn Asn Gln Tyr Ile Pro Lys Pro  
 355 360 365  
 Lys Pro Ala Arg Val Lys Lys Gly Arg Thr Arg His Ser  
 370 375 380

## (2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1785 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA  
(iii) HYPOTHETICAL: NO  
(iv) ANTI-SENSE: NO  
(vii) IMMEDIATE SOURCE:  
    (B) CLONE: HW038  
(ix) FEATURE:  
    (A) NAME/KEY: misc\_feature  
    (B) LOCATION: 1440..1714  
    (D) OTHER INFORMATION: /label= SAC\_24332

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GTGACGGTT GAGTCTGATG AGGAAGATCA AGAAAAGACT GAGGAGACTC CTTCCCAGCT	60
GGACCAGGAT ACTCAAGTAC AGGACATGGA TGAGGGTTCA GACGATGAAG AAGAAGGACA	120
GAAAGTGCCT CCACCCCCAG AGACACCCAT GCCTCCACCT CTGCCCCCAA CACCAGACCA	180
AGTCATTGTG CGAAAGGACT ATGATCCCAA AGCTTCCAAG CCCCTTCCTC CAGCTCCTGC	240
TCCTGATGAG TATCTTGTGT CCCCCATAAC TGGGGAGAAG ATTCCTGCCA GCAAATGCA	300
AGAACACATG CGCATCGGGC TTCTTGACCC CCGCTGGCTG GAGCAACGGG ATCGCTCCAT	360
CCGAGAGAAG CAGAGCANTG ATGAGGTGTA CGCACCAGGT CTGGATATTG AGAGCAGCCT	420
AAAGCAGTTG GCTGAGCGAC GTACTGACAT CTTTGGGGTA GAAGAGACAG CCATTGGTAA	480
GAAGATTGGT GAAGAGGAGA TCCAGAAGCC AGAGGAAAAG GTGACTTGGG ATGGCCATTC	540
AGGCAGCATG GCGCGGACCC AGCAGGCTGC CCAGGCCAAC ATCACCTCCA GGAGCAGATT	600
GAGGCCATCC ATAAGGCCAA GGGCCTGGTG CCAGAAGATG ACACCAAAGA GAAGATTGGC	660
CCTAGCAAGC CCAATGAAAT CCCTCAGCAG CCACCGCTCC ATCTTCAGCC ACCAATATCC	720
CCAGCTCAGC GCGCCCCATC ACGTCATTCC CCGGCCACCT GCGATGCCAC CTCCAGTCCG	780
CACCACTGTC GTATCTGCAG TGCCTGTCAT GCCTCGGCCC CCCATGGCAT CTGTGGTCCG	840
GCTGCCCCCA GGCTCAGTGA TTGCCCCGAT GCCACCCATC ATCCACGCAC CTAGGATCAA	900
CGTGGTACCC ATGCCTCCTG CAGCACCNCC CATCATGGCA CCCC GNCCAC CTCCCATGAT	960
TGTGCCAACA GCCTTTGTGC CTGCTCCCCC CTGTGGCTCC AGTGCCAGCC CCAGCTCCAA	1020
TGCCTCCCGT ACACCCCCCA CCTCCTATGG AAGATGAGCC TCCATCCAAG AAAC TAAAGA	1080
CCGAGGACAG CCTCATGCC T GAGGAAGAGT TCCTGCGCAG GAACAAGGGT CCAGTATCCA	1140
TCAAGGTGCA GGTACCCAAC ATGCAGGATA AGACAGAATG GAAGCTGAAT GGGCAGGGGC	1200
TGGTCTTCAC TCTCCCGCTC ACAGACCAGG TCTCTGTCAT CAAAGTGAAG ATTCATGAAG	1260
CNACGGGCAT GCCTGCAGGG AAGCAGAAAC TACAGTATGA AGGCATCTTC ATCAAGGACT	1320
CCAAC TCCCT GGCTTACTAC AACATGGCCA GTGGTGTCTGT CATCCACTTG GCCCTCAAGG	1380

```

AAAGAGGCGG CCGGAAGAAG TAGCCTCGAG AGATACTGCA AGTGGACTCC TGCCATTGTA      1440
CCTCTTTCCA CCACCAATTC CGGGAGTTTC CCTGGGACCT CGGGCACAGC CTGGGCATTT      1500
GTCTCCCTCT AAGGCAGTCT ATGGAATGGA TATCTGTTTA GATTGGATT GGAGGTCCTT      1560
CAAGGAAGCC CATTAGGAAT TGAAAACCCA GGGTTCTCCC ACAGGAGAGC TTCATAATAC      1620
CAAGTAATGT TTGGCTAATA GAGTATGGTT ACTACCATTC ATCATTGTGT ATCATGCTTT      1680
GGCCTCTAAT AAACCTCTTG AAGCCAGGCA TGGTACACAC ATGCCTGTAA TCCCGGGACT      1740
TGGGGGCGG GCAGAAGAAT TAAATGTTCA TCTTGGCTGT AAAAA                      1785

```

## (2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1054 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW039
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 231..617
  - (D) OTHER INFORMATION: /label= SAC\_24045

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```

TTTTTCAGAA TTGGCTGACT TTAATTTGGA AACAGCTTTT GGCTACAGAC ATTCAGGTTT      60
AAGGCACATT CCAAACCTACA AGGATCCCTG CCACTGGCAC GGGGCCTGGC TTTAATTCCTT      120
GCATGGTTTT ACATGGAAAG TGGAACAAAT GATTTTCTGG CCTCTTCTAA AAAAGAGTGT      180
GAGTTTGGCA AAAACAAAAC TCATGACACA TTCTCACCAG ACACCACAGT ACCGTTACAA      240
GACCTCCCTT TTGTCCTAGG CAAAGACGGG TGGCCCGCTC ACCTACCATT CTCTCCAGAA      300
ACATGTATAC AGTAAGCATG CTTGTATGCG TCTGTCTGGC AGCGATGGTT AAAAAAAGTT      360
TTCAAATAAA AATACAAAAA TCACACGTCT ATAACAAAGC TAAGAGTCTT AGCCACAAGC      420
AGCGAGGAAT GAGTGTAGGA ACCAGGAGAG ATAAACATCA ATTCTGCACA GCAAGGGGCT      480
GTGAGGGCCA GGGAGATGGT CACCTGAGAA CCTGGCCAC CTGGGCCCTT CGGCACAATA      540
CGTGGATGAG GGGGGACCAG AAGGAAAAGG TCATGGCTTT GGGGATGGGT GGATGGCTGG      600
GGGGTGCACT TGTGGGTACT AGATAGATTG GGGGAACTGG TGTGCTCTTC GGGACGTTTC      660

```

```

CTGGCCACTG CCAGCAGGGG GCCAGCAGGG GGCACAGGTG AGCCACTGCC ACCGCTCAGT      720
CGTTGCCACC ACAGATCTTC AGCAGGATCT TCCGATAATC CCCAGAAGTG TCTCCCGTGA      780
TATCGTGGTA CAGTGACTTG CCGTACATCC GCTTATACTC TGCTCGGATA TCCAGCAGGT      840
CAAGCTCGCT TCGAGACACC ATGATGCGAA TCAGGGTCCG GTCTTTTGTT CCTGCTCCCC      900
CTCATAGCCT TGTGAGCCT TTCGGCAAAG AAGGCAGGGG TGTTCTTGAG GCATTTACC      960
ACAGCCAGCA TGCCCTGCTC CAGGTCCCCA GACATCTCCG GCAGATGCTC TTCTCGATGT     1020
CTCGGCCTGT CATCCTCTGA TACTCGTTGA AACT                                1054

```

## (2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3306 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW040
- (ix) FEATURE:
  - (A) NAME/KEY: CDS
  - (B) LOCATION: 481..2433
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: complement (2395..2588)
  - (D) OTHER INFORMATION: /label= SAC\_24520

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

```

GTCGACGCCT GGTGTGGCTG GCGGCACCAG CTGAGTATTC CCTGAAAGAA TGAAGATGAC      60
TGGCCTAGAG GAGTCAGGAA TGTGACTGCT GGGAGATCCA GCTAGAGCAA CCATACCTGC     120
TTTTCCAGGA CCTTCCCCCT AACAGCTAAG GAGCCTTAGG TCTGGGAGAG GCAGGCACCT     180
GGCTGCAGAA GGGGAAACAA CCTTTTTCTC TTTCTCTCCT CCACTACTCT CCTGATGGCT     240
GCAGAAGAAG AAAGATCTAC CACGGGAAAC TCGGAGAGCT AGTCGGTTAA CTGAGCTGTT     300
GGATCTCCAC GGCACGAGGC AAATCCGTGA TGTGACTTTT TGTTTTTTCT CTAGGTCAG      360
CTGAGGGAAC AGCTCAGTTA CCTTAAGGGC GATAACTTTT TCAGGTTTAC TTGTGCCGAT     420
TGCTCGGCAG ACGGCAAGGA GCAGTCCGAG AGGCTGAAGC TGACATGGCA GCAAGTGGTC     480
ATG CTG GCT ATG TAC AAC TTG TCT CTG GAA GGA AGT GGA CGT CAA GGC      528
Met Leu Ala Met Tyr Asn Leu Ser Leu Glu Gly Ser Gly Arg Gln Gly
  1             5             10             15
TAC TTC CGG TGG AAA GAA GAT ATC TGT GCT TTT ATT GAG AAA CAT TGG      576

```

Tyr	Phe	Arg	Trp	Lys	Glu	Asp	Ile	Cys	Ala	Phe	Ile	Glu	Lys	His	Trp	
			20					25					30			
ACT	TTT	TTA	CTA	GGA	AAT	AGG	AAA	AAG	ACT	TCG	ACG	TGG	TGG	AGC	ACA	624
Thr	Phe	Leu	Leu	Gly	Asn	Arg	Lys	Lys	Thr	Ser	Thr	Trp	Trp	Ser	Thr	
		35					40					45				
GTA	GCA	GGT	TGC	CTC	AGC	GTG	GGA	AGT	CCT	GTT	TAT	TTC	CGT	TCA	GGT	672
Val	Ala	Gly	Cys	Leu	Ser	Val	Gly	Ser	Pro	Val	Tyr	Phe	Arg	Ser	Gly	
	50					55					60					
GCT	CAG	GAA	TTT	GGA	GAG	CCT	GGA	TGG	TGG	AAA	CTT	GTT	CAT	AAC	AGA	720
Ala	Gln	Glu	Phe	Gly	Glu	Pro	Gly	Trp	Trp	Lys	Leu	Val	His	Asn	Arg	
65					70					75					80	
CCC	CCA	ACA	ATG	AGA	CCA	GAG	GGA	GAG	AAG	CTG	TCT	GCC	TCC	ACT	TTG	768
Pro	Pro	Thr	Met	Arg	Pro	Glu	Gly	Glu	Lys	Leu	Ser	Ala	Ser	Thr	Leu	
			85						90					95		
AAA	GTG	AAA	GCT	TCA	AAA	CCA	ACA	CTG	GAT	CCC	ATC	ATT	ACC	GTT	GAA	816
Lys	Val	Lys	Ala	Ser	Lys	Pro	Thr	Leu	Asp	Pro	Ile	Ile	Thr	Val	Glu	
			100					105					110			
GGA	CTC	AGA	AAA	CGA	GCA	AGC	CGG	AAT	CCT	GTG	GAA	TCA	GCC	ATG	GAA	864
Gly	Leu	Arg	Lys	Arg	Ala	Ser	Arg	Asn	Pro	Val	Glu	Ser	Ala	Met	Glu	
		115					120					125				
TTG	AAA	GAG	AAG	CGG	TCT	CGA	ACA	CAG	GAA	GCT	AAA	GAC	ATC	AGA	AGA	912
Leu	Lys	Glu	Lys	Arg	Ser	Arg	Thr	Gln	Glu	Ala	Lys	Asp	Ile	Arg	Arg	
	130					135					140					
GCA	CAG	AAA	GAA	GCA	GCC	GGC	CTC	CTT	GAC	AGA	AGC	ACC	TCT	TCT	ACC	960
Ala	Gln	Lys	Glu	Ala	Ala	Gly	Leu	Leu	Asp	Arg	Ser	Thr	Ser	Ser	Thr	
145					150					155					160	
CCT	GTC	AAG	TTC	ATA	AGC	CGA	GGC	CGG	AGG	CCG	GAT	ATG	ATT	TTG	GAG	1008
Pro	Val	Lys	Phe	Ile	Ser	Arg	Gly	Arg	Arg	Pro	Asp	Met	Ile	Leu	Glu	
				165					170					175		
AAA	GGA	GAA	GTG	ATT	GAC	TTT	TCA	TCC	CTG	AGC	TCC	TCG	GAC	CGT	ACC	1056
Lys	Gly	Glu	Val	Ile	Asp	Phe	Ser	Ser	Leu	Ser	Ser	Ser	Asp	Arg	Thr	
			180					185					190			
CCC	CTC	ACA	AGC	CCC	TCT	CCT	TCT	CCC	TCT	CTG	GAT	TTC	TCT	GCC	CCA	1104
Pro	Leu	Thr	Ser	Pro	Ser	Pro	Ser	Pro	Ser	Leu	Asp	Phe	Ser	Ala	Pro	
		195					200					205				
GGG	ACA	CCA	GCG	TCA	CAT	TCT	GCC	ACG	CCT	AGC	TTG	CTC	TCT	GAA	GCA	1152
Gly	Thr	Pro	Ala	Ser	His	Ser	Ala	Thr	Pro	Ser	Leu	Leu	Ser	Glu	Ala	
	210					215					220					
GAC	CTC	ATC	CCA	GAC	GTG	ATG	CCA	CCC	CAA	GCT	CTA	TTT	CAT	GAT	GAT	1200
Asp	Leu	Ile	Pro	Asp	Val	Met	Pro	Pro	Gln	Ala	Leu	Phe	His	Asp	Asp	
225					230					235				240		
GAT	GAG	CTG	GAA	GGA	GAT	GGA	GTC	ATA	GAC	CCA	GGG	ATG	GAG	TAC	ATC	1248
Asp	Glu	Leu	Glu	Gly	Asp	Gly	Val	Ile	Asp	Pro	Gly	Met	Glu	Tyr	Ile	
				245					250					255		
CCA	CCC	CCA	GCT	GGG	TCA	GCT	TCT	GGG	CTG	ATG	GGG	AGC	AGA	AAG	AAG	1296
Pro	Pro	Pro	Ala	Gly	Ser	Ala	Ser	Gly	Leu	Met	Gly	Ser	Arg	Lys	Lys	
			260					265					270			

GTC AGA GCT GCA GAG CAG ATA AAA CAG GAA GTC GAC AGC GAG GAG GAG	1344
Val Arg Ala Ala Glu Gln Ile Lys Gln Glu Val Asp Ser Glu Glu Glu	
275 280 285	
AAG CCA GAC AGG ATG GAT GGA GAT AGC GAA GAC ACA GAT TCA AAC ACT	1392
Lys Pro Asp Arg Met Asp Gly Asp Ser Glu Asp Thr Asp Ser Asn Thr	
290 295 300	
TCT TTG CAC ACT AGA GCT CGA GAA AAG AGG AAG CCA CCC CTG GAG AAG	1440
Ser Leu His Thr Arg Ala Arg Glu Lys Arg Lys Pro Pro Leu Glu Lys	
305 310 315 320	
GAC ACG AAG CCC AAG GGG CCC AGG TAT ACG CCT GTG AGC ATC TAT GAG	1488
Asp Thr Lys Pro Lys Gly Pro Arg Tyr Thr Pro Val Ser Ile Tyr Glu	
325 330 335	
GAG AAG CTG CTG CTC AAG AGG CTG GAG GCA TGC CCT GGT GCC GTG GCC	1536
Glu Lys Leu Leu Leu Lys Arg Leu Glu Ala Cys Pro Gly Ala Val Ala	
340 345 350	
ATG ACT CCA GAA GCC CGG AGG CTG AAA CGG AAG CTG ATC GTC CGG CAA	1584
Met Thr Pro Glu Ala Arg Arg Leu Lys Arg Lys Leu Ile Val Arg Gln	
355 360 365	
GCA AAA AGG GAT CGG GGC TTA CCA CTG TTT GAC TTG GAC GAG GTT GTG	1632
Ala Lys Arg Asp Arg Gly Leu Pro Leu Phe Asp Leu Asp Glu Val Val	
370 375 380	
AAC GCA GCA CTT CTG TTA GTG GAT GGG ATT TAT GGA GCC AAA GAT GGA	1680
Asn Ala Ala Leu Leu Leu Val Asp Gly Ile Tyr Gly Ala Lys Asp Gly	
385 390 395 400	
GGA GCT TCC CGG CTT CCA GCT GGA CAA GCT ACA TAC CGG ACC ACC TGC	1728
Gly Ala Ser Arg Leu Pro Ala Gly Gln Ala Thr Tyr Arg Thr Thr Cys	
405 410 415	
CAG GAC TTC AGG ATC CTT GAC CGG TAC CAG ACT GCC TTA CCA GCC AGG	1776
Gln Asp Phe Arg Ile Leu Asp Arg Tyr Gln Thr Ala Leu Pro Ala Arg	
420 425 430	
AAA GGA TTC CGG CAC CAG ACC ACC AGA TTT TTG TAT CGT CTG GTG GGA	1824
Lys Gly Phe Arg His Gln Thr Thr Arg Phe Leu Tyr Arg Leu Val Gly	
435 440 445	
TCA GAA GAT CTG GCT GTG GAC CAA AGT ATT GTC AGC CCT TAC ACT TCT	1872
Ser Glu Asp Leu Ala Val Asp Gln Ser Ile Val Ser Pro Tyr Thr Ser	
450 455 460	
CGG ATC TTG AAA CCT TAT ATC AGG CGT GAT TAT GAG ACA AAG CCA CCC	1920
Arg Ile Leu Lys Pro Tyr Ile Arg Arg Asp Tyr Glu Thr Lys Pro Pro	
465 470 475 480	
AAA CTA CAG CTG CTG TCC CAG ATT CGC TCC CAC CTG CAC AAG AGT GAC	1968
Lys Leu Gln Leu Leu Ser Gln Ile Arg Ser His Leu His Lys Ser Asp	
485 490 495	
CCT CAC TGG ACA CCT GAA CCT GAT GCA CCT CTC GAT TAC TGC TAT GTC	2016
Pro His Trp Thr Pro Glu Pro Asp Ala Pro Leu Asp Tyr Cys Tyr Val	
500 505 510	
CGA CCA AAT CAC ATC CCA ACG ATC AAC TCC ATG TGT CAG GAA TTT TTC	2064
Arg Pro Asn His Ile Pro Thr Ile Asn Ser Met Cys Gln Glu Phe Phe	
515 520 525	

TGG CCT GGC ATT GAC CTG TCT GAG TGT CTG CAG TAT CCA GAC TTC AGT	2112
Trp Pro Gly Ile Asp Leu Ser Glu Cys Leu Gln Tyr Pro Asp Phe Ser	
530 535 540	
GTC GTA GTC CTT TAT AAA AAA GTC ATT GTT GCC TTT GGC TTC ATG GTT	2160
Val Val Val Leu Tyr Lys Lys Val Ile Val Ala Phe Gly Phe Met Val	
545 550 555 560	
CCC GAT GTG AAG TAC AAC GAA GCT TAC ATT TCA TTT CTG CTT GTT CAT	2208
Pro Asp Val Lys Tyr Asn Glu Ala Tyr Ile Ser Phe Leu Leu Val His	
565 570 575	
CCC GAG TGG AGG AGA GCA GGG ATT GCC ACG TTC ATG ATC TAT CAT CTG	2256
Pro Glu Trp Arg Arg Ala Gly Ile Ala Thr Phe Met Ile Tyr His Leu	
580 585 590	
ATT CAG ACG TGC ATG GGC AAG GAT GTG ACT CTT CAC GTC TCG GCG AGC	2304
Ile Gln Thr Cys Met Gly Lys Asp Val Thr Leu His Val Ser Ala Ser	
595 600 605	
AAC CCT GCC ATG CTG CTG TAC CAG AAG TTT GGC TTC AAG ACA GAG GAG	2352
Asn Pro Ala Met Leu Leu Tyr Gln Lys Phe Gly Phe Lys Thr Glu Glu	
610 615 620	
TAT GTT TTG GAT TTT TAT GAT AAG TAT TAC CCA CTG GAG AGT ACA GAG	2400
Tyr Val Leu Asp Phe Tyr Asp Lys Tyr Tyr Pro Leu Glu Ser Thr Glu	
625 630 635 640	
TGT AAA CAT GCA TTC TTC CTG AGG CTC CGA CGC TGAGGTGAAC TCAGTTCTTC	2453
Cys Lys His Ala Phe Leu Arg Leu Arg	
645 650	
CCAGCAAACA GCAGATGCTG TCAGAGAGTG GCCTCAGAAG TCGTGCAGGC CCTTCTCCTC	2513
GAAGTGGGCC CTGGCCTCCT GTGAGTAATG GCCTCCACAG GGTCCAGCTG GGTGAGTGAA	2573
GCCATGTGCA GAGGTACAAG TGGCTGTTCA ATCCAGAAAC CATCCAGCAG GCCCTGCGTG	2633
GCCTTCCAGC AGTATCTTGG ACTCCACTGT GGACCAGAAC CAAGTGACTC TGCCATTCTT	2693
TTTCCACACA ATGAGTGGGC ACAAAGAGC TTCTGAGCTG AAGCATTGGA GAATGGGCCCT	2753
GCTCCTCCTA TGTGAGAGCT CAGGAAGTAT GAAAGATGTT AAAGACTGCC AGGGCCCTCA	2813
CTGTCTGGG CCGAGAGGTT ATGAGGTGTC AGGAATGACT CCAACCCTGT CTCAGCCTGT	2873
GCTGTGCTGT CATCGAACCC CAGAGGATTC ATGGGGCACT GAAGTCCTTT TGTCTCTTAC	2933
TGACTTTTCGT CCTGGTCAGG AAAAAGATAC TCGGTGTCTG TTTCCAGACT GCTGCTGCTT	2993
TACTTATCTA CCTGTCAGCA CCCCTTGCAG GGACATTGAC CTGGTACAAG AAGTTTACTC	3053
TTGGCATGGC CTGGAAGTA GCGGAGAGG AACGGTTGAG CTTGATACAT TAGATGCCTT	3113
CCGAGGCCCT TCTCATGCTC AGGACAGTTG CACCACAGAA GTGGAGGTGT GGAATGGCGC	3173
AGGCAGGCCT GGGCGCCCTG TCCACCTCAT GAAGCTGCTT CCTGGGTTGC ACATGTTGGA	3233
GTTGGACTAC AGGTATCTAA GTGCATTCCA AAAGCATCAG ATATCTCCAA TAAAAATCCT	3293
TACCATACAA AAA	3306

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 651 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```

Met Leu Ala Met Tyr Asn Leu Ser Leu Glu Gly Ser Gly Arg Gln Gly
 1           5           10           15
Tyr Phe Arg Trp Lys Glu Asp Ile Cys Ala Phe Ile Glu Lys His Trp
 20           25           30
Thr Phe Leu Leu Gly Asn Arg Lys Lys Thr Ser Thr Trp Trp Ser Thr
 35           40           45
Val Ala Gly Cys Leu Ser Val Gly Ser Pro Val Tyr Phe Arg Ser Gly
 50           55           60
Ala Gln Glu Phe Gly Glu Pro Gly Trp Trp Lys Leu Val His Asn Arg
 65           70           75           80
Pro Pro Thr Met Arg Pro Glu Gly Glu Lys Leu Ser Ala Ser Thr Leu
 85           90           95
Lys Val Lys Ala Ser Lys Pro Thr Leu Asp Pro Ile Ile Thr Val Glu
100           105           110
Gly Leu Arg Lys Arg Ala Ser Arg Asn Pro Val Glu Ser Ala Met Glu
115           120           125
Leu Lys Glu Lys Arg Ser Arg Thr Gln Glu Ala Lys Asp Ile Arg Arg
130           135           140
Ala Gln Lys Glu Ala Ala Gly Leu Leu Asp Arg Ser Thr Ser Ser Thr
145           150           155           160
Pro Val Lys Phe Ile Ser Arg Gly Arg Arg Pro Asp Met Ile Leu Glu
165           170           175
Lys Gly Glu Val Ile Asp Phe Ser Ser Leu Ser Ser Ser Asp Arg Thr
180           185           190
Pro Leu Thr Ser Pro Ser Pro Ser Pro Ser Leu Asp Phe Ser Ala Pro
195           200           205
Gly Thr Pro Ala Ser His Ser Ala Thr Pro Ser Leu Leu Ser Glu Ala
210           215           220
Asp Leu Ile Pro Asp Val Met Pro Pro Gln Ala Leu Phe His Asp Asp
225           230           235           240
Asp Glu Leu Glu Gly Asp Gly Val Ile Asp Pro Gly Met Glu Tyr Ile
245           250           255
Pro Pro Pro Ala Gly Ser Ala Ser Gly Leu Met Gly Ser Arg Lys Lys
260           265           270

```



Val Arg Ala Ala Glu Gln Ile Lys Gln Glu Val Asp Ser Glu Glu Glu  
 275 280 285  
 Lys Pro Asp Arg Met Asp Gly Asp Ser Glu Asp Thr Asp Ser Asn Thr  
 290 295 300  
 Ser Leu His Thr Arg Ala Arg Glu Lys Arg Lys Pro Pro Leu Glu Lys  
 305 310 315 320  
 Asp Thr Lys Pro Lys Gly Pro Arg Tyr Thr Pro Val Ser Ile Tyr Glu  
 325 330 335  
 Glu Lys Leu Leu Leu Lys Arg Leu Glu Ala Cys Pro Gly Ala Val Ala  
 340 345 350  
 Met Thr Pro Glu Ala Arg Arg Leu Lys Arg Lys Leu Ile Val Arg Gln  
 355 360 365  
 Ala Lys Arg Asp Arg Gly Leu Pro Leu Phe Asp Leu Asp Glu Val Val  
 370 375 380  
 Asn Ala Ala Leu Leu Leu Val Asp Gly Ile Tyr Gly Ala Lys Asp Gly  
 385 390 395 400  
 Gly Ala Ser Arg Leu Pro Ala Gly Gln Ala Thr Tyr Arg Thr Thr Cys  
 405 410 415  
 Gln Asp Phe Arg Ile Leu Asp Arg Tyr Gln Thr Ala Leu Pro Ala Arg  
 420 425 430  
 Lys Gly Phe Arg His Gln Thr Thr Arg Phe Leu Tyr Arg Leu Val Gly  
 435 440 445  
 Ser Glu Asp Leu Ala Val Asp Gln Ser Ile Val Ser Pro Tyr Thr Ser  
 450 455 460  
 Arg Ile Leu Lys Pro Tyr Ile Arg Arg Asp Tyr Glu Thr Lys Pro Pro  
 465 470 475 480  
 Lys Leu Gln Leu Leu Ser Gln Ile Arg Ser His Leu His Lys Ser Asp  
 485 490 495  
 Pro His Trp Thr Pro Glu Pro Asp Ala Pro Leu Asp Tyr Cys Tyr Val  
 500 505 510  
 Arg Pro Asn His Ile Pro Thr Ile Asn Ser Met Cys Gln Glu Phe Phe  
 515 520 525  
 Trp Pro Gly Ile Asp Leu Ser Glu Cys Leu Gln Tyr Pro Asp Phe Ser  
 530 535 540  
 Val Val Val Leu Tyr Lys Lys Val Ile Val Ala Phe Gly Phe Met Val  
 545 550 555 560  
 Pro Asp Val Lys Tyr Asn Glu Ala Tyr Ile Ser Phe Leu Leu Val His  
 565 570 575  
 Pro Glu Trp Arg Arg Ala Gly Ile Ala Thr Phe Met Ile Tyr His Leu  
 580 585 590  
 Ile Gln Thr Cys Met Gly Lys Asp Val Thr Leu His Val Ser Ala Ser  
 595 600 605

Asn Pro Ala Met Leu Leu Tyr Gln Lys Phe Gly Phe Lys Thr Glu Glu  
610 615 620

Tyr Val Leu Asp Phe Tyr Asp Lys Tyr Tyr Pro Leu Glu Ser Thr Glu  
625 630 635 640

Cys Lys His Ala Phe Phe Leu Arg Leu Arg Arg  
645 650

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4234 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW041

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 3407..3482
- (D) OTHER INFORMATION: /label= SAC\_24317

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTGACCCCTC CTTCTGCCCT GGCTCCTGTT GATGAGCACA GTGCCGTCCC TGATCAGTTC	60
TTTCTTCCTG TGAATTTGCT TCGTAGCCCA TCAGGCTGGT AGACTGGCAG CACGGTCAGT	120
GTCACACGAG GGACAGGACT TCCCTGAAGT GAAGCAGCAA TGTCATGGAA GCAAACCCCT	180
GTGAGCGAGA GACAGGCCCC GAGGCTGCTG CCACAAGGTC TCATGTTTCAT TGCTGGGGAC	240
ACTAGCTTCC TACAAACCCC AGGAGACGCT GCCCATAGCA CGGGTGGCCC TTCAGGCTCC	300
TGAACTATCT ATCCCCAGAT CAGTAGGATT TGTTTTCCTAA GTCCCCCTCC TCATTCCTTC	360
TTTGCTCAAG AGAAAAAGA AAAGTGGTAA CTCTGTGACT CTAAGTCCCT CTTAAATGCA	420
GCGCTGAATC TAGCTTGGCT CGGAGATCTC TGACTTGTCC CTAGGGTGAG ACTGTGGCCT	480
CCCCAGGTTG AGGTGGCAGC TGTATTGTGA GAGATTAGCC TGGAGAGCAG AGAAGGGGGG	540
AGGGGGGATA GGGAGGAGAG AGCAATCGTG TCTCCCTACA GGACCCTGGG CCATCTGACA	600
AGACAGAGGA GGTGGGACTG GCTTCCCCAT GCTGCACATG GCAGTGGGAA CCAGGCTTGG	660
AGTCCCCCAC CCCCTCCTTA CCTCTTCTCT CAGACCTTGT AGCCCTAGCT TGCCCTGTGC	720
TCTGCAGGGA AGGGGATGAG AGGACAGCAT AAGAAGTCTT GCCAGTTGAG GTGGGGCTGG	780
AAGATGAAGC AGGTGGGTGT AGGAATCATT TTGGATAGGG CATGTTGTTT GGTGACGGC	840
AACTAATTCT CTTAAGAGCA TATTGTCTCC AGTTAATGTT GGAGACAACC CAGAAAGGGA	900

AGGGTATAGA GATGGGTAAT TAATGAGATA TTCCAGCCCT GGAGCAATGT CTGAGCCACC	960
CCAAACTTAG GCTGTCTCCC AGCATAGGAG TGGGGCACTG CCCCTTACTC AAATTCTAGT	1020
CTAAGTGTTT CAGACACATT GTACACTTGG ACAGGAAATC TCCTGTTCCA GCATACCCCTC	1080
TCCCCCATACT CCTCTCCCCC ATACCCCTC CCCCATACCT TCTCCCCCAT ACTGTTAGTA	1140
CGTTTCAGTT GTTCAAAAGC AAAGGGGGCG CAGACTAGGC TTAACTGCA TGTGGAGTGT	1200
AGATTTTGCC TAAGAATCCA TGTGAAGAA GAGAGAGGCT ATGTCTTTAA GGTCGAGGCC	1260
CTTCAGTCTT CTGTCTTGT AATTCTCTTG AAGCGTGAGC CCCAGGCACT CAGAAGCCAT	1320
GTTGTACTCC GGAAGTTCA TGGGACCTGA ACTCCTTCAA CACCGGGAAC AACAGTACTG	1380
CATAGGTA CTGGTACTGA CTGTGCAAGC ACTAGAAAA GTACCCGTG AATGGCCCTC	1440
ACAGCATCAG GGAACGAGG CCGAGTTCAC TCTCTGATT GCTCAGATGC AGAAGCTGGA	1500
CCCGTGAGGT GTACCAGCCC ACCTCCAAGC TTGACACAGA GGCTAGCACT TAAAGCAAAG	1560
ACTCAGAAGA GGAGTCTCAG TCACAGGCTG CACAGCCTTG AACATGATAT CTAGAACTAT	1620
GCCTGGGTGA TCGGTGGCCC AACTTGTGAC ACTTGTGTGC ACATAGGGAA GCTGACGCTC	1680
CTCTCGGGGC CTGAAAACAG GACTAAACCC AAACCCATA TAACTAACCA TACCCTGCTC	1740
ATCCACAGGC GTGGATGGAA GTCTTCACGG TCTGTTACAA GCACAAGTAC AAAGCCGATA	1800
AAAATAAGTC AAAGCCCTC CTCGTGGTGC AGAGGGCTCC TCCTCAGATC TCTTGGAGCC	1860
TTTGATCATA GAGCTTGACG GTCTCTCTC GGGATAAAG ACAGGTAAAG GCTGTGGCCT	1920
TCAGCGGTGA CTGCCCCACA CCCTGCTTTC CTTAGGGGTG GAGACAAGGA ATGGTGGGGC	1980
CAGTGTGTT CAGAGAGAAG TCTCTCTTTG GAGCCCATGT ATGGGGCTTG ACCTTTTGCT	2040
GTGAGTGCAT TTAGAGGCAA GTGAATTCTG GCCTCAGTAC AGAGAACCCA GCGGTTTCAG	2100
CACTCAAGAA AGAAGCCGCA GTGTCCCTA CAGAGAGGCT CTGTCTGCT TGGCCCCAGC	2160
CACCTGCGCC ACACATTTAA TAAAGTGGCT GAAACTCCTC CTTTCCAAGA GCTAGTGTGC	2220
TGGGTCTGT CCTCCCCACA GCCTTCTGAG TCCCAGAGAC TTGTGTGAGC AGTCGGATAA	2280
GGCCAATCTG GGAGACAGGG GATGACCATG ACACAAGGTG TTTCCAGGCA GTCCCTTGT	2340
GCGTAGGCTG CAGAGAGCTG TCCTGACTGG GAGCTTGGGT TGTACGAACA CATGGTGCCT	2400
TCTTTGTCTG TGCTGAACGA GTAGGGTGTG TTTCTCTGTG TTCTCTGAGT AAGATCACAC	2460
CCTCTTTAAG CAAAGGGGTG TTGTGGATAT TAGAAGCAAG ATTGATAAAT GTTCTTAAAT	2520
GGTGGTGCTA GAGAAAAGTC AGTCCCCGTA ACTTCCTAAA AGCTGAGTCA GCCAGTCTCT	2580
AAGTATGTAC ATGTTGGCTT TCTTCTATTT TTTTCTAGT TAATCAGTGT GTTTTAAATG	2640
TTTTTAGTTG ATGTGAGGAT TTTTAAAG GCATTTGTGT GTGGTGTGGC TCCATGGGAG	2700
ATAGGCTGTG TAATAAGAAC CCTCAGAAGC CAGCCAGTCA TCCAAGGGGC GCCCAAGTTC	2760
AGTGGATTTC TTCAGGAATT GGAGAGCAGT ATTTTCCCT CTTTTCAGTA GTATGCCTGT	2820

TCTCATTTAG TAATTAAATG GTATATTGCA TTATATGTGA ATGCATACAT AAACGGTGGA	2880
ATTTCTCTTA ATATCCAAAA AAGTCCTTCT GTTGCCCTCT GAGCTCTCTC CTTGGGGAGG	2940
ACTGCAGGGC TCACTGTGTA AAATGGTTGT GGACTCCTCT AGCAGAATGG CTTCCAAGGC	3000
CTGGCCCAT TCTCAGATTC CCTTTGAAGT GAGAGGTCAT TGTCTTTTGA GAATGGATTT	3060
AAGTTTTTGA AATACCCAC GCTGTTTACA GTGCAAGCGA CTGAACACAG AACAATTAGT	3120
CAGAATCAGA ATCCCAGGGC CATCAAGATG ACTCCACGGA AAAGGGACTG CCACCAACCA	3180
GAGCCAAGAC CTGAGTTTGG TCCCCAGGAA CCCCATGGCA GCAGGAGAGA GTCAGCCGAT	3240
GCAAGTGTCA CACACACACA TAAACACACA CATACACATG CAACACATAC ACACAAACAC	3300
ATACATACAC ACACACAAAC ACATACACAC ATGCACACAC ACACACACAC ACACACACAC	3360
AGAGTTTTTG GAAGAATATC TATGATGTAG GAAAAGGTTT GTTAGTACCT GCAACCTCTA	3420
GGGCCTCAT TCCAAGCCTG CTGCTTCTCT TGCCAGGACT TATACGTGTG CAACCCCTGC	3480
GTACACTACA GACCTAAGTA CATACTCGCT AACATGAACA CCCAGTAGGT ATTCTGTAAA	3540
ACGGATATAC TTAAGTGCAT GTTATTGTCG TGTGTGCCGC GACTGAGCTG TGTCTTTGTC	3600
ATGAACTTT GCTTCTTTGC CAAATTAAGA CTCTGCCCTC CCTCCGAGGA AGGAAATATA	3660
CAGCAGTTTC TGGAATCAGC TTCTCAAGTT TTCCCGCCTG TGTGTGTGTT ACTTGTGGCT	3720
TTGCCTGAGA GTCTGTCTC ACGTACACAC CTGTGAGTAT TTATTACAAG GCAGAACGGG	3780
GACAGGCGTA GCCATAGTAC TTGTGTTTGA ATGTTTCCAG GGAGGGACAA TGGAGCCCCC	3840
TCCTCAGTGA GGTGCTCACA GAATCTGTGT CAAATTCTGT GAACGGGGTC CAACACTAGT	3900
ACCAGCTTCA CCCCCAACCCT GAGTGCCCTG GAGCCACTGC CCTCTCTACT GCCTTGTGTG	3960
TGCATGGGTG TGTTGTGTG TGTCTCTGTT GGGGCATGTG CCGCATTTCT CAGTGGACTT	4020
TGGTATCCAC ATGTAACCTT TGGCATCTGT GTTCTCTGGAC CGCATTGACA GATCTGTGGT	4080
GTGCAGATCT GTGTAGATGG GTCTGTGTGC AAAGTGGGGA GTCCCCTGAA AGTGACTATG	4140
GATATTGTTG GTCAGCCAAA GACTTTCCTA TTCTTTGCTG CTTAACTTT GTGCCTTAAT	4200
ATTGTAAATA TAAACGGTTA AAACAAAAGT CGAC	4234

## (2) INFORMATION FOR SEQ ID NO:29:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1792 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW042

## (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 8..889

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (1409..1618)  
(D) OTHER INFORMATION: /label= SAC\_24017

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GTCTGACC	AAG	GAT	GGT	GGA	AGT	TTC	AGT	GGT	ACT	GGT	GGC	CTT	CCA	GAT	49	
Lys	Asp	Gly	Gly	Ser	Phe	Ser	Gly	Thr	Gly	Gly	Leu	Pro	Asp			
1				5						10						
TAT	TCT	GCC	CCC	AAT	CCC	ATT	AAA	GTG	ACC	CAT	CGG	TGC	TAC	ATC	CTT	97
Tyr	Ser	Ala	Pro	Asn	Pro	Ile	Lys	Val	Thr	His	Arg	Cys	Tyr	Ile	Leu	
15				20						25					30	
GAG	AAT	GAC	ACA	GTC	CAG	TGT	GAC	TTG	GAC	CTG	TAC	AAG	TCC	CTG	CAG	145
Glu	Asn	Asp	Thr	Val	Gln	Cys	Asp	Leu	Asp	Leu	Tyr	Lys	Ser	Leu	Gln	
				35				40						45		
GCT	TGG	AAA	GAC	CAC	AAA	TTG	CAC	ATC	GAC	CAT	GAG	ATT	GAA	ACC	CTG	193
Ala	Trp	Lys	Asp	His	Lys	Leu	His	Ile	Asp	His	Glu	Ile	Glu	Thr	Leu	
			50				55					60				
CAG	AAC	AAA	ATT	AAG	AAC	CTT	CGA	GAA	GTC	AGG	GGT	CAC	CTG	AAG	AAG	241
Gln	Asn	Lys	Ile	Lys	Asn	Leu	Arg	Glu	Val	Arg	Gly	His	Leu	Lys	Lys	
	65					70					75					
AAG	CGG	CCA	GAA	GAA	TGT	GAC	TGT	CAC	AAA	GTC	AGT	TAC	CAC	AGC	CAA	289
Lys	Arg	Pro	Glu	Glu	Cys	Asp	Cys	His	Lys	Val	Ser	Tyr	His	Ser	Gln	
	80				85						90					
CAC	AAA	GGC	CGC	CTC	AAG	CAC	AAG	GGC	TCC	AGC	CTG	CAC	CCT	TTC	AGG	337
His	Lys	Gly	Arg	Leu	Lys	His	Lys	Gly	Ser	Ser	Leu	His	Pro	Phe	Arg	
	95			100				105						110		
AAG	GGT	CTG	CAG	GAG	AAA	GAC	AAG	GTG	TGG	CTG	TTG	CGG	GAG	CAG	AAA	385
Lys	Gly	Leu	Gln	Glu	Lys	Asp	Lys	Val	Trp	Leu	Leu	Arg	Glu	Gln	Lys	
			115					120						125		
CGC	AAG	AAG	AAA	CTG	CGC	AAG	CTG	CTG	AAG	CGA	CTG	CAG	AAC	AAT	GAC	433
Arg	Lys	Lys	Lys	Leu	Arg	Lys	Leu	Leu	Lys	Arg	Leu	Gln	Asn	Asn	Asp	
			130				135						140			
ACG	TGC	AGC	ATG	CCG	GGC	CTC	ACG	TGC	TTT	ACT	CAC	GAC	AAC	CAC	CAC	481
Thr	Cys	Ser	Met	Pro	Gly	Leu	Thr	Cys	Phe	Thr	His	Asp	Asn	His	His	
	145					150					155					
TGG	CAG	ACC	GCG	CCG	CTC	TGG	ACG	TTG	GGA	CCC	TTC	TGC	GCC	TGC	ACC	529
Trp	Gln	Thr	Ala	Pro	Leu	Trp	Thr	Leu	Gly	Pro	Phe	Cys	Ala	Cys	Thr	
	160				165					170						
AGC	GCC	AAC	AAT	AAC	ACG	TAC	TGG	TGC	CTG	CGG	ACC	ATC	AAC	GAG	ACC	577
Ser	Ala	Asn	Asn	Asn	Thr	Tyr	Trp	Cys	Leu	Arg	Thr	Ile	Asn	Glu	Thr	
	175			180				185						190		
CAC	AAC	TTC	CTC	TTC	TGT	GAA	TTC	GCA	ACT	GGC	TTC	ATA	GAA	TAC	TTT	625

His	Asn	Phe	Leu	Phe	Cys	Glu	Phe	Ala	Thr	Gly	Phe	Ile	Glu	Tyr	Phe		
			195						200					205			
GAC	CTC	AGT	ACG	GAC	CCC	TAC	CAG	CTG	ATG	AAC	GCG	GTC	AAC	ACA	CTG		673
Asp	L	u	Ser	Thr	Asp	Pro	Tyr	Gln	Leu	Met	Asn	Ala	Val	Asn	Thr	Leu	
			210					215					220				
GAC	AGG	GAC	GTT	CTC	AAC	CAG	CTG	CAC	GTC	CAG	CTC	ATG	GAG	CTA	AGG		721
Asp	Arg	Asp	Val	Leu	Asn	Gln	Leu	His	Val	Gln	Leu	Met	Glu	Leu	Arg		
		225					230					235					
AGC	TGC	AAA	GGC	TAC	AAG	CAG	TGT	AAC	CCC	CGG	ACC	CGC	AAT	ATG	GAC		769
Ser	Cys	Lys	Gly	Tyr	Lys	Gln	Cys	Asn	Pro	Arg	Thr	Arg	Asn	Met	Asp		
		240				245					250						
CTG	GGG	CTT	AGA	GAC	GGA	GGA	AGC	TAT	GAG	CAA	TAC	AGG	CAG	TTT	CAG		817
Leu	Gly	Leu	Arg	Asp	Gly	Gly	Ser	Tyr	Glu	Gln	Tyr	Arg	Gln	Phe	Gln		
	255				260					265					270		
CGT	CGA	AAA	TGG	CCA	GAA	ATG	AAG	AGA	CCC	TCT	TCC	AAA	TCA	CTG	GGA		865
Arg	Arg	Lys	Trp	Pro	Glu	Met	Lys	Arg	Pro	Ser	Ser	Lys	Ser	Leu	Gly		
			275					280						285			
CAG	CTA	TGG	GAA	GGT	TGG	GAA	GGC	TAATCGGCCA	TGGAGAGGAA	CCTCCAAAAC							919
Gln	Leu	Trp	Glu	Gly	Trp	Glu	Gly										
			290														
CAGAGGCCTC	CTGTGGCTGC	CCAGGCGTGC	AAAAATCACC	CAATTCCGAG	GAGATCGATG												979
CTGGAAGTGG	GAGACTTGAC	AGGAGGCAGG	GTCTGCACTT	GGGACAGGAA	ATCCCGAGGA												1039
GAACGCCAAG	ACTTCCAGAG	GCTCATTTGT	TTGCCCGCT	TTGCTTCGGA	TTGAACCTCA												1099
CCGGCCGCAC	AAGGATGCGT	CCTGACACCT	GGAGTCTCCG	CTCACCCCTTT	CAGAGGCTCA												1159
CCAAGACAAA	GGAACATAAT	CCCATGGACT	TTCCTCCAGA	GATGGAAAAT	GCTGGGATTT												1219
ACCCATCCCC	TCCCCTGTAC	CCCCTTCCCC	ACTAGGGAAA	CAAGCTTGTT	TTAACCTTCT												1279
TATTCTTTGG	AGAAAGCACG	GATGTCCCAG	GTGCTGTCAA	CAGCGCAGTC	TTGACAATGG												1339
TCTATAGCAC	AAAACAGTAC	CATTACCTG	GCTGGATGAG	CCGGCTGCCC	CAGAAGCTGC												1399
CTTCACTGTA	CATATGTGAC	TGCTCACATG	TAACCAACAC	AGGGACTTGT	AGGGGGAATC												1459
TCACTAATAC	GAAATCCCAT	TTTCAAGAGT	CGCGGTGTCA	ATAAACACTC	TGTGGCTGGT												1519
GTCAAGGGTA	ATCCCTTGAG	CTTTCAGACA	TTTCTGTTC	TGCCCAGGAT	TCGTTCCCTTT												1579
GTTATCCATA	TCCATCCCAG	AACTGATGTT	TTCTAAGGTA	CTGAAACCCC	AAGTTGATGT												1639
GTGTCTGTG	TTTAAATGAC	ATTGTATTTG	TAAAGCAATT	TTGTAGTATA	AAAGTACCAT												1699
CTTCCAGTGT	TCACATCCCC	CCCAGCCAAT	GACTAGCTCT	TGGTATGAAA	AAAAAAATCT												1759
TTGAATTTTT	GTAAAAA	AAAAAAGTC	GAC														1792

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 294 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Lys Asp Gly Gly Ser Phe Ser Gly Thr Gly Gly L u Pro Asp Tyr Ser  
 1 5 10 15  
 Ala Pro Asn Pro Ile Lys Val Thr His Arg Cys Tyr Ile Leu Glu Asn  
 20 25 30  
 Asp Thr Val Gln Cys Asp Leu Asp Leu Tyr Lys Ser Leu Gln Ala Trp  
 35 40 45  
 Lys Asp His Lys Leu His Ile Asp His Glu Ile Glu Thr Leu Gln Asn  
 50 55 60  
 Lys Ile Lys Asn Leu Arg Glu Val Arg Gly His Leu Lys Lys Lys Arg  
 65 70 75 80  
 Pro Glu Glu Cys Asp Cys His Lys Val Ser Tyr His Ser Gln His Lys  
 85 90 95  
 Gly Arg Leu Lys His Lys Gly Ser Ser Leu His Pro Phe Arg Lys Gly  
 100 105 110  
 Leu Gln Glu Lys Asp Lys Val Trp Leu Leu Arg Glu Gln Lys Arg Lys  
 115 120 125  
 Lys Lys Leu Arg Lys Leu Leu Lys Arg Leu Gln Asn Asn Asp Thr Cys  
 130 135 140  
 Ser Met Pro Gly Leu Thr Cys Phe Thr His Asp Asn His His Trp Gln  
 145 150 155 160  
 Thr Ala Pro Leu Trp Thr Leu Gly Pro Phe Cys Ala Cys Thr Ser Ala  
 165 170 175  
 Asn Asn Asn Thr Tyr Trp Cys Leu Arg Thr Ile Asn Glu Thr His Asn  
 180 185 190  
 Phe Leu Phe Cys Glu Phe Ala Thr Gly Phe Ile Glu Tyr Phe Asp Leu  
 195 200 205  
 Ser Thr Asp Pro Tyr Gln Leu Met Asn Ala Val Asn Thr Leu Asp Arg  
 210 215 220  
 Asp Val Leu Asn Gln Leu His Val Gln Leu Met Glu Leu Arg Ser Cys  
 225 230 235 240  
 Lys Gly Tyr Lys Gln Cys Asn Pro Arg Thr Arg Asn Met Asp Leu Gly  
 245 250 255  
 Leu Arg Asp Gly Gly Ser Tyr Glu Gln Tyr Arg Gln Phe Gln Arg Arg  
 260 265 270  
 Lys Trp Pro Glu Met Lys Arg Pro Ser Ser Lys Ser Leu Gly Gln Leu  
 275 280 285  
 Trp Glu Gly Trp Glu Gly  
 290

## (2) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2261 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW043

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 8..634

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 582..808
- (D) OTHER INFORMATION: /label= SAC\_24533

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTCGACG	GCT	CCA	CCA	GGT	TCA	AGA	CAC	AAA	CCT	GTG	CCT	ACC	AAG	TCA	49	
Ala	Pro	Pro	Gly	Ser	Arg	His	Lys	Pro	Val	Pro	Thr	Lys	Ser			
1				5						10						
ACA	CCA	GCT	GGC	ACA	GAA	CTA	GGT	AGC	ACA	GGA	AGC	GAC	GGG	GGA	GAT	97
Thr	Pro	Ala	Gly	Thr	Glu	Leu	Gly	Ser	Thr	Gly	Ser	Asp	Gly	Gly	Asp	
15				20				25					30			
GCA	GTA	GAA	TAC	AAC	CCC	AAC	CTC	CTG	GAT	GAC	CCA	CAG	TGG	CCG	TGT	145
Ala	Val	Glu	Tyr	Asn	Pro	Asn	Leu	Leu	Asp	Asp	Pro	Gln	Trp	Pro	Cys	
				35				40					45			
GGG	AAG	CAC	AAG	CGT	GTA	CTC	ATC	TTT	GCC	TCA	TAC	ATG	ACC	ACT	GTT	193
Gly	Lys	His	Lys	Arg	Val	Leu	Ile	Phe	Ala	Ser	Tyr	Met	Thr	Thr	Val	
			50				55					60				
ATA	GAG	TAT	GTG	AAG	CCT	GCA	GAC	CTC	AAA	AAG	GAC	ATG	AAC	GAG	ACC	241
Ile	Glu	Tyr	Val	Lys	Pro	Ala	Asp	Leu	Lys	Lys	Asp	Met	Asn	Glu	Thr	
	65					70					75					
TTC	AGG	GAG	AAG	TTC	CCT	CAT	ATC	AAA	CTG	ACA	TTG	AGC	AAA	ATT	AGG	289
Phe	Arg	Glu	Lys	Phe	Pro	His	Ile	Lys	Leu	Thr	Leu	Ser	Lys	Ile	Arg	
	80					85					90					
AGT	TTA	AAA	CGG	GAG	ATG	CGG	AAC	CTT	TCT	GAA	GAG	TGC	AGC	TTG	GAG	337
Ser	Leu	Lys	Arg	Glu	Met	Arg	Asn	Leu	Ser	Glu	Glu	Cys	Ser	Leu	Glu	
	95			100						105				110		
CCT	GTG	ACT	GTG	TCC	ATG	GCC	TAT	GTG	TAC	TTT	GAG	AAG	CTT	GCA	CTG	385
Pro	Val	Thr	Val	Ser	Met	Ala	Tyr	Val	Tyr	Phe	Glu	Lys	Leu	Ala	Leu	
			115					120					125			
CAG	GGC	AAG	CTC	AAC	AAA	CAG	AAC	CGC	AAA	CTG	TGT	GCT	GGC	GCC	TGT	433
Gln	Gly	Lys	Leu	Asn	Lys	Gln	Asn	Arg	Lys	Leu	Cys	Ala	Gly	Ala	Cys	
			130					135					140			



GTT CTG TTG GCT GCC AAG ATC AGC AGT GAC CTC CGC AAG AGT GAA GTG Val Leu Leu Ala Ala Lys Ile Ser Ser Asp Leu Arg Lys Ser Glu Val 145 150 155	481
AAG CAG CTT ATT GAC AAG CTG GAG GAA AGG TTC CGG TTC AAC AGG AAG Lys Gln Leu Ile Asp Lys Leu Glu Glu Arg Phe Arg Phe Asn Arg Lys 160 165 170	529
GAT CTC ATT GGA TTT GAG TTC ACG GTG CTT GTG GCT TTG GAA CTG GCC Asp Leu Ile Gly Phe Glu Phe Thr Val Leu Val Ala Leu Glu Leu Ala 175 180 185 190	577
CTG TAC CTC CCG GAG AAC CAA GTA TTA CCT CAC TAC AGA CGC CTC ACC Leu Tyr Leu Pro Glu Asn Gln Val Leu Pro His Tyr Arg Arg Leu Thr 195 200 205	625
CAG CAG TTC TAGCCCAGAA CAGTGCACCC CACAGGGAGC CTGCTTTGGC Gln Gln Phe	674
AGGAAGGCGA GTAGCTTGGC TAGAGCACTG ACTAGTAGTG TCCCTGTCCT TCCTCTAGTA	734
ACTGCACTTG CCCCTCAGTC CAGGGCTGCA CTTAGCTTCC CTTGGAAAGA GGCAGTGCTC	794
TTATTTTGGA GTGTACCCCA TCTTCACCTT GAGGCCCTGT TGAGGGTTTT CCTGAAGCAA	854
ATGAACATCG TGGGGCTTGG TTTGCTTTTC ATGTGCCTGA GACCAGCCTG GTTTATGTTG	914
TCAATCCTGA GGTCCAGCTT TCATTTTGCC TAGCCATGAA ACCATTCCCTG TAGAGACCTA	974
GGTGTCCAGA CCTTCTGAAC ATACTGGAAT TGCCAGGCTT CTGTCTCCTG GTGCCTCCAC	1034
CCAGGCAGAG CTCCCCATC AGCACTTTTT TTTGGTCAAG ATGCAGGAAA GGGCAGTCTG	1094
CTCACACTGG ACCTTGTTAA AGCATCTGCC ACTGTGAAAC CCCTGTGCCA CCACATACAG	1154
TGTTTGGCCA ACACCACCAC CCACCCCAAC CCCTGCCTCC TGCCAGCACT TATCCCTGCG	1214
GTACTGGACC ATTTAAGACA ACCGCTGACT GCGGGCTCTT GACTTTCCTG AGCAATATCA	1274
TGATGGAGAA AAGGTATTAA CTTGAGATTT GGTTCCTATA ACAGAAAAGA CGGCCCATTA	1334
AGGCCTGCTG CTGACTCTGC ACCACACTTG ATCCTCAAGC AGGAAGGTAC TGACCCTGCC	1394
ACCTAACAGC ACATCAGTGA CTCCTCACAA ATAGATGCCA TCCAGGCCAA CTTGCCTGTG	1454
TATTGACTGT GCCACTGTCT TCCTACTGGC ACCAGACAAG AGCCATCCTC CACCTTCCCA	1514
GGCCTGTTAT TCCCTACAAC TGGTCACCAG CCTGGCAGAT GCCAGTCCCA CTGCTCTCAA	1574
GTGCTTGTTT CCACAGCCAC TCATCCCAGA AATGCTTACC CTCTCAAGCC TGACACCACT	1634
CAACTACCTA ATGGATGTTA GTAGCTTAAA CAGAGACCTT GAAGGATAGT AGAAGGATAG	1694
TAAACTCTAG CCTTCTACCC TGCCCATGCT TATGCTGGTG TAAGCACGCT ATCCTTAGAG	1754
CACCTGGGTG CCATCTGCAG ACAGTGATGT CTTACTTAAA ATGAACCTTA GATACAGCTG	1814
ACTGGTCGGC TATCCCTGGG GGTGATACCC TGGTAGCTTA TTCTAGTTAT GTTAAAGCTG	1874
GCAACAGTGA TAATTTAAGC TTCTATATTT ATGATCTTAC CAAGCTGGTG TGCTGTTAGT	1934

CCTTCCCTCC TTGCCCTCAG TAATGCTAAG AGCTGCTGAC ATGTGGTATG ATAAGCACTG 1994  
 TTACAACAAA ACCAACCTTT AAACCAAATG ACAGGCTAAA CACAGTGTCT GTCCTAGACT 2054  
 AAGTCCAGAA ATCCGAGGTG TGCCCACGGT GTCCAGCTGG ACTGTGACCC AATGCCACAG 2114  
 ACCTGGTGCT GTACCTTCAG CACTCAACAG GCACAGCTGT GTGTGTGACC TCAGAATGCT 2174  
 GCTTATGTAT GTCCAGCCGG TTCTGAATGG AAGGTGTAA ATGTTTTTGT TAAAAGTAA 2234  
 ACTTGTTTTT GCAAAAAAAA AGTCGAC 2261

## (2) INFORMATION FOR SEQ ID NO:32:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 209 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Ala Pro Pro Gly Ser Arg His Lys Pro Val Pro Thr Lys Ser Thr Pro  
 1 5 10 15  
 Ala Gly Thr Glu Leu Gly Ser Thr Gly Ser Asp Gly Gly Asp Ala Val  
 20 25 30  
 Glu Tyr Asn Pro Asn Leu Leu Asp Asp Pro Gln Trp Pro Cys Gly Lys  
 35 40 45  
 His Lys Arg Val Leu Ile Phe Ala Ser Tyr Met Thr Thr Val Ile Glu  
 50 55 60  
 Tyr Val Lys Pro Ala Asp Leu Lys Lys Asp Met Asn Glu Thr Phe Arg  
 65 70 75 80  
 Glu Lys Phe Pro His Ile Lys Leu Thr Leu Ser Lys Ile Arg Ser Leu  
 85 90 95  
 Lys Arg Glu Met Arg Asn Leu Ser Glu Glu Cys Ser Leu Glu Pro Val  
 100 105 110  
 Thr Val Ser Met Ala Tyr Val Tyr Phe Glu Lys Leu Ala Leu Gln Gly  
 115 120 125  
 Lys Leu Asn Lys Gln Asn Arg Lys Leu Cys Ala Gly Ala Cys Val Leu  
 130 135 140  
 Leu Ala Ala Lys Ile Ser Ser Asp Leu Arg Lys Ser Glu Val Lys Gln  
 145 150 155 160  
 Leu Ile Asp Lys Leu Glu Glu Arg Phe Arg Phe Asn Arg Lys Asp Leu  
 165 170 175  
 Ile Gly Phe Glu Phe Thr Val Leu Val Ala Leu Glu Leu Ala Leu Tyr  
 180 185 190  
 Leu Pro Glu Asn Gln Val Leu Pro His Tyr Arg Arg Leu Thr Gln Gln  
 195 200 205

Phe

## (2) INFORMATION FOR SEQ ID NO:33:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1538 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW044

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 24..1100

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1161..1378)
- (D) OTHER INFORMATION: /label= SAC\_24216

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

GTCTGACCTCC ATAGAGTTCC GGC ATG AAG CCA CCG CAG AGA CGG CGG AAG	50
Met Lys Pro Pro Gln Arg Arg Arg Lys	
1 5	
GTC CCC GCA CGC TAC ACA GGT GAA GCT ACC GGC CCC ACA GCT TGG AGC	98
Val Pro Ala Arg Tyr Thr Gly Glu Ala Thr Gly Pro Thr Ala Trp Ser	
10 15 20 25	
CCC CGC GAG ATG CGG CAC CTA CTG AGA CTA CTC CAG GCT CGG CGC GGC	146
Pro Arg Glu Met Arg His Leu Leu Arg Leu Leu Gln Ala Arg Arg Gly	
30 35 40	
CAG CCG GAG CCA GAC GCC AAG GAG CTG GCC AAG GAG CTG CGA GGC CGG	194
Gln Pro Glu Pro Asp Ala Lys Glu Leu Ala Lys Glu Leu Arg Gly Arg	
45 50 55	
AGC GAG GCC GAG ATC TGT CGC TTT ATC CAG CAG CTC AAA GGC CGG GTG	242
Ser Glu Ala Glu Ile Cys Arg Phe Ile Gln Gln Leu Lys Gly Arg Val	
60 65 70	
GTT CGA GAG GCT ATT CGG AAG GTG CAA CCA GGT GGC AGA GAT GGT CCA	290
Val Arg Glu Ala Ile Arg Lys Val Gln Pro Gly Gly Arg Asp Gly Pro	
75 80 85	
AGG CAT CGA GAG ACA CAG CTC CCA GCC CCC ATA GAG GTA TGG ATG GAT	338
Arg His Arg Glu Thr Gln Leu Pro Ala Pro Ile Glu Val Trp Met Asp	
90 95 100 105	
CTT GCT GAG AAA TTA ACA GGC CCA CTG GAG GAA GCC CTG ACT GCA GCT	386
Leu Ala Glu Lys Leu Thr Gly Pro Leu Glu Glu Ala Leu Thr Ala Ala	
110 115 120	

TTC TCC CAG GTA CTC ACC ATT GCT GCT GCG GAA CCT CTC AGC CTC CTG	434
Phe Ser Gln Val Leu Thr Ile Ala Ala Ala Glu Pro Leu Ser Leu Leu	
125 130 135	
CAT TCC AAG CCA GGC AAG CCC ACA AAG GCC TGT GGA AGG GCA CTG GTG	482
His Ser Lys Pro Gly Lys Pro Thr Lys Ala Cys Gly Arg Ala Leu Val	
140 145 150	
TTC TTG AGC ACC CAA GAT GCG CAG AAG GAT TCT GCC CCT GAA GGT TCT	530
Phe Leu Ser Thr Gln Asp Ala Gln Lys Asp Ser Ala Pro Glu Gly Ser	
155 160 165	
GGG CCT GAA CCT ATG ACT GCT GCT GAC CCC ACT CCT GAG GCC TCT GTC	578
Gly Pro Glu Pro Met Thr Ala Ala Asp Pro Thr Pro Glu Ala Ser Val	
170 175 180 185	
CCT GAA CCT AAG GCC TCC GGT ACC AAC CCT GAG ACC TCT GGC CTT GTC	626
Pro Glu Pro Lys Ala Ser Gly Thr Asn Pro Glu Thr Ser Gly Leu Val	
190 195 200	
CCA GAG GTC TCT GTC CCT GAC CCT GAT GCA CCA ACT GAG TCC CTG GCT	674
Pro Glu Val Ser Val Pro Asp Pro Asp Ala Pro Thr Glu Ser Leu Ala	
205 210 215	
GGA TCC TCC ACA GAG AGA GAC TTT GCC GTG GAC TTT GAG AAG ATC TAC	722
Gly Ser Thr Thr Glu Arg Asp Phe Ala Val Asp Phe Glu Lys Ile Tyr	
220 225 230	
AAA TAC CTG TCC TCC TCC TCC AGA GGT GGC CAT GGC CCT GAG CTC TCA	770
Lys Tyr Leu Ser Ser Ser Ser Arg Gly Gly His Gly Pro Glu Leu Ser	
235 240 245	
GCA GTT GAG TCC GCC GTG GTC CTT AAC CTG CTC ATG TCA CTT GCA GAG	818
Ala Val Glu Ser Ala Val Val Leu Asn Leu Leu Met Ser Leu Ala Glu	
250 255 260 265	
GAA CTG CCC CAC CTG CCT TGC ACA GCC CTG GTG GAC CAT CTG ACT AAA	866
Glu Leu Pro His Leu Pro Cys Thr Ala Leu Val Asp His Leu Thr Lys	
270 275 280	
ACA TAT GCT CAA CTG ACG GCT CCC CAG GCC TCT CTC GAT AGA GAA AAG	914
Thr Tyr Ala Gln Leu Thr Ala Pro Gln Ala Ser Leu Asp Arg Glu Lys	
285 290 295	
AGG CCC AGG CCG GGG ACT GAA GAT GGA GGG ACT GAC TCC ACG GGG CCA	962
Arg Pro Arg Pro Gly Thr Glu Asp Gly Gly Thr Asp Ser Thr Gly Pro	
300 305 310	
GAA GAG ACT GGC CAA GGC AGT CCT CAA GCT TCT GAG CCC ACA GAA CCG	1010
Glu Glu Thr Gly Gln Gly Ser Pro Gln Ala Ser Glu Pro Thr Glu Pro	
315 320 325	
AGG CTG ACC TGG AAA GCA GCT GGG GTC TGC CCA CTG AAC CCA TTC CTG	1058
Arg Leu Thr Trp Lys Ala Ala Gly Val Cys Pro Leu Asn Pro Phe Leu	
330 335 340 345	
GTG CCC CTG GAC CTC CTG AGC CAG GCC CTC GCC CCT GCC AGG	1100
Val Pro Leu Asp Leu Leu Ser Gln Ala Leu Ala Pro Ala Arg	
350 355	
TGAAGGATGC AGCAGGCTGG CTATGCTGGT TTCTTTCTAC ATCGTCTGTC CAGCACTGGT	1160
ACTGGTGTGCG GGAACGTGAT TGTGGCTTCC TGACAGTGTC CGGCCATAGG CAAGCCCTTC	1220

AGACACTGCA TGCCTAGTA GCCTGGAAAG TGAAGCACAG TTGAAGGAAG GGGTTTAACT 1280  
 ATTGGGGATT TGAGACTGCA CAAAGGCGGA ATTCCTGGGC TGGTAGATTG CGGCAAGTGT 1340  
 CCCTGGCCCC GCTGTGCCTG GTTCTCTGCC ATAAGAGTAC CACATCATCA TCCCTTTGTG 1400  
 CATGGCACTC GAACCCACCC CCAGCACACC CTCCCCACCT CAACTGACCA CTGTACCATC 1460  
 CTCTTCGCCC TTCTCTGGGT GTCCTGTGAC CCTCAATAAA GCTGTCATCT AACTAGCTGT 1520  
 GAAAAAAAAA AAGTCGAC 1538

## (2) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 359 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Lys Pro Pro Gln Arg Arg Arg Lys Val Pro Ala Arg Tyr Thr Gly  
 1 5 10 15  
 Glu Ala Thr Gly Pro Thr Ala Trp Ser Pro Arg Glu Met Arg His Leu  
 20 25 30  
 Leu Arg Leu Leu Gln Ala Arg Arg Gly Gln Pro Glu Pro Asp Ala Lys  
 35 40 45  
 Glu Leu Ala Lys Glu Leu Arg Gly Arg Ser Glu Ala Glu Ile Cys Arg  
 50 55 60  
 Phe Ile Gln Gln Leu Lys Gly Arg Val Val Arg Glu Ala Ile Arg Lys  
 65 70 75 80  
 Val Gln Pro Gly Gly Arg Asp Gly Pro Arg His Arg Glu Thr Gln Leu  
 85 90 95  
 Pro Ala Pro Ile Glu Val Trp Met Asp Leu Ala Glu Lys Leu Thr Gly  
 100 105 110  
 Pro Leu Glu Glu Ala Leu Thr Ala Ala Phe Ser Gln Val Leu Thr Ile  
 115 120 125  
 Ala Ala Ala Glu Pro Leu Ser Leu Leu His Ser Lys Pro Gly Lys Pro  
 130 135 140  
 Thr Lys Ala Cys Gly Arg Ala Leu Val Phe Leu Ser Thr Gln Asp Ala  
 145 150 155 160  
 Gln Lys Asp Ser Ala Pro Glu Gly Ser Gly Pro Glu Pro Met Thr Ala  
 165 170 175  
 Ala Asp Pro Thr Pro Glu Ala Ser Val Pro Glu Pro Lys Ala Ser Gly  
 180 185 190  
 Thr Asn Pro Glu Thr Ser Gly Leu Val Pro Glu Val Ser Val Pro Asp  
 195 200 205

Pro Asp Ala Pro Thr Glu Ser Leu Ala Gly Ser Ser Thr Glu Arg Asp  
 210 215 220  
 Phe Ala Val Asp Phe Glu Lys Ile Tyr Lys Tyr Leu Ser Ser Ser Ser  
 225 230 235 240  
 Arg Gly Gly His Gly Pro Glu Leu Ser Ala Val Glu Ser Ala Val Val  
 245 250 255  
 Leu Asn Leu Leu Met Ser Leu Ala Glu Glu Leu Pro His Leu Pro Cys  
 260 265 270  
 Thr Ala Leu Val Asp His Leu Thr Lys Thr Tyr Ala Gln Leu Thr Ala  
 275 280 285  
 Pro Gln Ala Ser Leu Asp Arg Glu Lys Arg Pro Arg Pro Gly Thr Glu  
 290 295 300  
 Asp Gly Gly Thr Asp Ser Thr Gly Pro Glu Glu Thr Gly Gln Gly Ser  
 305 310 315 320  
 Pro Gln Ala Ser Glu Pro Thr Glu Pro Arg Leu Thr Trp Lys Ala Ala  
 325 330 335  
 Gly Val Cys Pro Leu Asn Pro Phe Leu Val Pro Leu Asp Leu Leu Ser  
 340 345 350  
 Gln Ala Leu Ala Pro Ala Arg  
 355

## (2) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2277 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: HW045

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 9..1628

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1973..2134
- (D) OTHER INFORMATION: /label= SAC\_24028

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GTCGACCA GGG CGT GAT GGC TCC TAC TAC CAG CTC TTT GTT CAT GGT GGC  
 Gly Arg Asp Gly Ser Tyr Tyr Gln Leu Phe Val His Gly Gly  
 1 5 10

50

CGC CTC CAG CCA GTC CTA AGG CAG AAA GCC TGC CGA GGC ATG AAC TGG Arg Leu Gln Pro Val Leu Arg Gln Lys Ala Cys Arg Gly Met Asn Trp 15 20 25 30	98
ATA GCT GGG CTT CGG ATG GTG CCC GAT GGC AGT ATG GTC GTC TTG GGT Ile Ala Gly Leu Arg Met Val Pro Asp Gly Ser Met Val Val Leu Gly 35 40 45	146
TTC CAT GCC AAC GAG TTT GTA GTG TGG AGC CCG CGG TCC CAT GAG AAG Phe His Ala Asn Glu Phe Val Val Trp Ser Pro Arg Ser His Glu Lys 50 55 60	194
CTG CAC ATC GTC AAC TGC GGG GGA GGG CAC CGC TCC TGG GCC TTT TCT Leu His Ile Val Asn Cys Gly Gly His Arg Ser Trp Ala Phe Ser 65 70 75	242
GAT ACT GAG GCA GCC ATG GCC TTT GCC TAC CTT AAG GAT GGT GAT GTC Asp Thr Glu Ala Ala Met Ala Phe Ala Tyr Leu Lys Asp Gly Asp Val 80 85 90	290
ATG CTC TAC CGG GCT CTA GGA GGC TGC ATC AGG CCA AAC GTG ATT CTC Met Leu Tyr Arg Ala Leu Gly Gly Cys Ile Arg Pro Asn Val Ile Leu 95 100 105 110	338
CGG GAG GGT CTG CAT GGC CGG GAA ATC ACA TGT GTA AAG CGT GTG GGC Arg Glu Gly Leu His Gly Arg Glu Ile Thr Cys Val Lys Arg Val Gly 115 120 125	386
ACT GTT ACC CTG GGC CCT GAA TTT GAG GTA CCC AAC TTG GAG CAT CCT Thr Val Thr Leu Gly Pro Glu Phe Glu Val Pro Asn Leu Glu His Pro 130 135 140	434
GAC TCC CTG GAG CCT GGC AGT GAG GGG CCT GGT CTG ATT GAC ATC GTG Asp Ser Leu Glu Pro Gly Ser Glu Gly Pro Gly Leu Ile Asp Ile Val 145 150 155	482
ATA ACA GGC AGT GAG GAC ACT ACT GTC TGT GTG CTA GCA CTT CCC ACC Ile Thr Gly Ser Glu Asp Thr Thr Val Cys Val Leu Ala Leu Pro Thr 160 165 170	530
ACC ACA GGC GCA GCC CAC GCC CTC ACT GCT GTC TGT AAC CAT ATC TCC Thr Thr Gly Ala Ala His Ala Leu Thr Ala Val Cys Asn His Ile Ser 175 180 185 190	578
TCC GTG CGA GCC CTG GCA GTG TGG GGT GTT GGC ACC CCA GGT GGC CCA Ser Val Arg Ala Leu Ala Val Trp Gly Val Gly Thr Pro Gly Gly Pro 195 200 205	626
CAG GAT TCT CAC CCA GGG CTC ACG GCT CAG GTA GTG TCT GCA GGG GGT Gln Asp Ser His Pro Gly Leu Thr Ala Gln Val Val Ser Ala Gly Gly 210 215 220	674
CGA GCC GAG ATG CAC TGC TTC AGC CTC ATG ATC ACT CCG GAT GCC AGC Arg Ala Glu Met His Cys Phe Ser Leu Met Ile Thr Pro Asp Ala Ser 225 230 235	722
ACC CCA AGC CGC CTT GCC TGC CAC ATC ATG CAC CTT TCA TCC CAC CGG Thr Pro Ser Arg Leu Ala Cys His Ile Met His Leu Ser Ser His Arg 240 245 250	770
CTA GAT GAG TAC TGG GAC CGG CAG CGC AAC CGG CAC CGG ATG ATC AAG Leu Asp Glu Tyr Trp Asp Arg Gln Arg Asn Arg His Arg Met Ile Lys 255 260 265 270	818

GTG GAC CCT GAG ACC AGG TAC ATG TCT CTT GCC ATT TGT GAG CTT GAC Val Asp Pro Glu Thr Arg Tyr Met Ser Leu Ala Ile Cys Glu Leu Asp 275 280 285	866
TCC GAT AGG CCT GGC CTT GGC CCC GGC CCC CTT GTG GCT GCA GCC TGT Ser Asp Arg Pro Gly Leu Gly Pro Gly Pro Leu Val Ala Ala Ala Cys 290 295 300	914
AGT GAC GGG GCA GTG AGG CTT TTT CTC TTG CAG GAC TCT GGG CGA ATT Ser Asp Gly Ala Val Arg Leu Phe Leu Leu Gln Asp Ser Gly Arg Ile 305 310 315	962
CTG CAT CTC CTT GCT GAG ACT TTC CAC CAT AAG CGG TGT GTC CTC AAG Leu His Leu Leu Ala Glu Thr Phe His His Lys Arg Cys Val Leu Lys 320 325 330	1010
GTC CAT TCG TTC ACA CAT GAG GCA CCC AAC CAG CGT CGG AGG CTG ATC Val His Ser Phe Thr His Glu Ala Pro Asn Gln Arg Arg Arg Leu Ile 335 340 345 350	1058
CTG TGC AGT GCA GCT ACA GAT GGC AGC ATA GCC TTC TGG GAT CTC ACC Leu Cys Ser Ala Ala Thr Asp Gly Ser Ile Ala Phe Trp Asp Leu Thr 355 360 365	1106
ACG GCA ATG GAC CGA GGC TCT ACT ACC CTG GAG CCT CCA GCA CAC CCT Thr Ala Met Asp Arg Gly Ser Thr Thr Leu Glu Pro Pro Ala His Pro 370 375 380	1154
GGG CTT CCC TAC CAG ATG GGC ACC CCC TGC CTG ACC GTC CAG GCC CAT Gly Leu Pro Tyr Gln Met Gly Thr Pro Cys Leu Thr Val Gln Ala His 385 390 395	1202
AGC TGT GGT GTC AAC AGC CTG CAC ACT TTG CCT ACA CCT GAG GGC CAT Ser Cys Gly Val Asn Ser Leu His Thr Leu Pro Thr Pro Glu Gly His 400 405 410	1250
CAT CTT GTG GCC AGT GGC AGT GAG GAC GGG TCC CTG CAT GTC TTC ACG His Leu Val Ala Ser Gly Ser Glu Asp Gly Ser Leu His Val Phe Thr 415 420 425 430	1298
CTT GCT GTG AAG ATG CCA GAG CTG GAA GAG GCT GAT GGG GAG GCT GAG Leu Ala Val Lys Met Pro Glu Leu Glu Glu Ala Asp Gly Glu Ala Glu 435 440 445	1346
TTG GTG CCC CAG TTA TGT GTC CTA GAT GAA TAT TCT GTC CCC TGT GCA Leu Val Pro Gln Leu Cys Val Leu Asp Glu Tyr Ser Val Pro Cys Ala 450 455 460	1394
CAT GCT GCC CAT GTG ACG GGC ATC AAG ATC CTA AGT CCC AAG CTG ATG His Ala Ala His Val Thr Gly Ile Lys Ile Leu Ser Pro Lys Leu Met 465 470 475	1442
GTC TCA GCC TCC ATA GAC CAA CGG CTG ACC TTC TGG CGT CTT GGA AAT Val Ser Ala Ser Ile Asp Gln Arg Leu Thr Phe Trp Arg Leu Gly Asn 480 485 490	1490
GGT GAG CCC ACC TTC ATG AAT AGC ACT GTG TAC CAC GTG CCA GAT GTG Gly Glu Pro Thr Phe Met Asn Ser Thr Val Tyr His Val Pro Asp Val 495 500 505 510	1538
GCT GAC ATG GAC TGC TGG CCT GTG AAC CCT GAG TTT GGC CAC CGC TGT Ala Asp Met Asp Cys Trp Pro Val Asn Pro Glu Phe Gly His Arg Cys	1586



515	520	525	
GCT CTT GCG GGT CAG GGA CTC GAG GTT TAC AAC TGG TAT GAC			1628
Ala Leu Ala Gly Gln Gly Leu Glu Val Tyr Asn Trp Tyr Asp			
530	535	540	
TGAGTTACCA CAGTAGTTGG AGGACTGAGC CTGGGGCTTG ACCACAGACA GCGGAGCAGG			1688
GATCAGCTGT CTGTGTCATG CTCTATGTGC TCTGAGGAGG TGAGGCAGTA CCATGGGTTC			1748
CTGTAGGGTG TTGCATAGGT GGACCAAGAA TATGCCTCAC TCTTCACAAT AGGATGAAAC			1808
TGTATTTATT CTGACTTTAA GTGCCCAACA TCTGTGAGGT TTCTGTGAGG TCTTGTTTT			1868
TTCCCAAGTTG ATGCTTTTAT AAACATTCCC AGCTATTGGG CCCTTAGATG TGGCTCAGCG			1928
GAGGGAGGCC CAGCATGGCC AAGCCTGTGT GGAACACCTC ACGTACTGCC CTCAAAGCT			1988
GTAGGCGAGC AAACATCTGA CCAAAGAGGT GTGGCCGAGG TTCCTGAAAG AAAAGCAACC			2048
AGGCCCATCC TCATTTCCCG AGCCTGAGCC CTTGTCCATA TTTCCACAG ACCCTCCCTT			2108
GCTGTATGCT CACCCCTAGA ATGTGTACGC GGTATAGTA TGAGCTGAAA TCCATGCTGA			2168
GCTGCACCAG GAACTTGCAC ACCTAGAGAC AGAGGTTGGA TCGTTGAGCT GTTGTCTTT			2228
TTCTTGTGTC ACAACCCAGA ATAAAGAATA GTGTGTGTAG TGTCAAAA			2277

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 540 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Gly Arg Asp Gly Ser Tyr Tyr Gln Leu Phe Val His Gly Gly Arg Leu			
1	5	10	15
Gln Pro Val Leu Arg Gln Lys Ala Cys Arg Gly Met Asn Trp Ile Ala			
20	25	30	
Gly Leu Arg Met Val Pro Asp Gly Ser Met Val Val Leu Gly Phe His			
35	40	45	
Ala Asn Glu Phe Val Val Trp Ser Pro Arg Ser His Glu Lys Leu His			
50	55	60	
Ile Val Asn Cys Gly Gly Gly His Arg Ser Trp Ala Phe Ser Asp Thr			
65	70	75	80
Glu Ala Ala Met Ala Phe Ala Tyr Leu Lys Asp Gly Asp Val Met Leu			
85	90	95	
Tyr Arg Ala Leu Gly Gly Cys Ile Arg Pro Asn Val Ile Leu Arg Glu			
100	105	110	
Gly Leu His Gly Arg Glu Ile Thr Cys Val Lys Arg Val Gly Thr Val			
115	120	125	

Thr Leu Gly Pro Glu Phe Glu Val Pro Asn Leu Glu His Pro Asp Ser  
130 135 140

Leu Glu Pro Gly Ser Glu Gly Pro Gly Leu Ile Asp Ile Val Ile Thr  
145 150 155 160

Gly Ser Glu Asp Thr Thr Val Cys Val Leu Ala Leu Pro Thr Thr Thr  
165 170 175

Gly Ala Ala His Ala Leu Thr Ala Val Cys Asn His Ile Ser Ser Val  
180 185 190

Arg Ala Leu Ala Val Trp Gly Val Gly Thr Pro Gly Gly Pro Gln Asp  
195 200 205

Ser His Pro Gly Leu Thr Ala Gln Val Val Ser Ala Gly Gly Arg Ala  
210 215 220

Glu Met His Cys Phe Ser Leu Met Ile Thr Pro Asp Ala Ser Thr Pro  
225 230 235 240

Ser Arg Leu Ala Cys His Ile Met His Leu Ser Ser His Arg Leu Asp  
245 250 255

Glu Tyr Trp Asp Arg Gln Arg Asn Arg His Arg Met Ile Lys Val Asp  
260 265 270

Pro Glu Thr Arg Tyr Met Ser Leu Ala Ile Cys Glu Leu Asp Ser Asp  
275 280 285

Arg Pro Gly Leu Gly Pro Gly Pro Leu Val Ala Ala Ala Cys Ser Asp  
290 295 300

Gly Ala Val Arg Leu Phe Leu Leu Gln Asp Ser Gly Arg Ile Leu His  
305 310 315 320

Leu Leu Ala Glu Thr Phe His His Lys Arg Cys Val Leu Lys Val His  
325 330 335

Ser Phe Thr His Glu Ala Pro Asn Gln Arg Arg Arg Leu Ile Leu Cys  
340 345 350

Ser Ala Ala Thr Asp Gly Ser Ile Ala Phe Trp Asp Leu Thr Thr Ala  
355 360 365

Met Asp Arg Gly Ser Thr Thr Leu Glu Pro Pro Ala His Pro Gly Leu  
370 375 380

Pro Tyr Gln Met Gly Thr Pro Cys Leu Thr Val Gln Ala His Ser Cys  
385 390 395 400

Gly Val Asn Ser Leu His Thr Leu Pro Thr Pro Glu Gly His His Leu  
405 410 415

Val Ala Ser Gly Ser Glu Asp Gly Ser Leu His Val Phe Thr Leu Ala  
420 425 430

Val Lys Met Pro Glu Leu Glu Glu Ala Asp Gly Glu Ala Glu Leu Val  
435 440 445

Pro Gln Leu Cys Val Leu Asp Glu Tyr Ser Val Pro Cys Ala His Ala  
450 455 460

Ala His Val Thr Gly Ile Lys Ile Leu Ser Pro Lys Leu Met Val S r  
465 470 475 480

Ala Ser Ile Asp Gln Arg Leu Thr Phe Trp Arg Leu Gly Asn Gly Glu  
485 490 495

Pro Thr Phe Met Asn Ser Thr Val Tyr His Val Pro Asp Val Ala Asp  
500 505 510

Met Asp Cys Trp Pro Val Asn Pro Glu Phe Gly His Arg Cys Ala Leu  
515 520 525

Ala Gly Gln Gly Leu Glu Val Tyr Asn Trp Tyr Asp  
530 535 540

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: HW046

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 9..1160

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1523..1760)
- (D) OTHER INFORMATION: /label= SAC\_24036

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GTCGACGT GAC TTC CTT GTC AGG TTC ACC CCC GCT GCA GCA GGG CTC TGC	50
Asp Phe Leu Val Arg Phe Thr Pro Ala Ala Ala Gly Leu Cys	
1 5 10	
TCT CTG ACT GAG TGG TGC TGT GTT ACA TGG ATC TGT CAT TTC TAT TCA	98
Ser Leu Thr Glu Trp Cys Cys Val Thr Trp Ile Cys His Phe Tyr Ser	
15 20 25 30	
GTC ATC AGT TGG CAT TGC AGT AAC ACC TAT GGC AAT GAA GAA AGA GAC	146
Val Ile Ser Trp His Cys Ser Asn Thr Tyr Gly Asn Glu Glu Arg Asp	
35 40 45	
AGA ATT TGG GAT TTT CAT CTA AGG AAA ATG ATA AAT AAC AAC CAG ACG	194
Arg Ile Trp Asp Phe His Leu Arg Lys Met Ile Asn Asn Asn Gln Thr	
50 55 60	
TGT GCT GCA GGG CAG GAC TCC GTG CCC TAC GTG ACC TGT ATG CTT CAC	242
Cys Ala Ala Gly Gln Asp Ser Val Pro Tyr Val Thr Cys Met Leu His	
65 70 75	

GTG CTG GAA GAG TGG TTG GGT GTG GAA CAG TTG GAG GAC TAC CTG AAT Val Leu Glu Glu Trp Leu Gly Val Glu Gln Leu Glu Asp Tyr Leu Asn 80 85 90	290
TTT GCG AAC CAT CTC TTG TGG GTC TTC ACC CCG CTG ATC CTT TTA ATA Phe Ala Asn His Leu Leu Trp Val Phe Thr Pro Leu Ile Leu Leu Ile 95 100 105 110	338
CTT CCG TAC TTT ACC ATC TTC CTT CTC TAC CTT ACT ATT ATT TTC CTC Leu Pro Tyr Phe Thr Ile Phe Leu Leu Tyr Leu Thr Ile Ile Phe Leu 115 120 125	386
CAC ATC TAT AAG AGG AAG AAT GTG TTA AAA GAA GCC TAC TCT CAC AAC His Ile Tyr Lys Arg Lys Asn Val Leu Lys Glu Ala Tyr Ser His Asn 130 135 140	434
TTG TGG GAT GGC GCG AGG AAA ACA GTG GCA ACC CTG TGG GAT GGA CAT Leu Trp Asp Gly Ala Arg Lys Thr Val Ala Thr Leu Trp Asp Gly His 145 150 155	482
GCG GCG GTT TGG CAT GGT TAT GAA GTT CAC GGG ATG GAA AAG ATA CCA Ala Ala Val Trp His Gly Tyr Glu Val His Gly Met Glu Lys Ile Pro 160 165 170	530
GAA GGA CCA GCA CTT ATA ATT TTT TAT CAT GGA GCT ATT CCC ATA GAC Glu Gly Pro Ala Leu Ile Ile Phe Tyr His Gly Ala Ile Pro Ile Asp 175 180 185 190	578
TTT TAC TAC TTC ATG GCT AAA ATT TTC ATC CAC AAA GGC AGA ACT TGC Phe Tyr Tyr Phe Met Ala Lys Ile Phe Ile His Lys Gly Arg Thr Cys 195 200 205	626
CGA GTG GTA GCT GAC CAC TTT GTC TTT AAA ATC CCA GGG TTC AGT TTA Arg Val Val Ala Asp His Phe Val Phe Lys Ile Pro Gly Phe Ser Leu 210 215 220	674
TTA CTT GAT GTA TTT TGT GCT CTT CAT GGA CCA AGA GAA AAA TGT GTT Leu Leu Asp Val Phe Cys Ala Leu His Gly Pro Arg Glu Lys Cys Val 225 230 235	722
GAA ATC TTG AGG AGT GGT CAC TTG TTA GCT ATT TCA CCG GGT GGA GTT Glu Ile Leu Arg Ser Gly His Leu Leu Ala Ile Ser Pro Gly Gly Val 240 245 250	770
CGA GAA GCC TTA CTT AGT GAT GAA ACC TAC AAC ATC ATA TGG GGT AAT Arg Glu Ala Leu Leu Ser Asp Glu Thr Tyr Asn Ile Ile Trp Gly Asn 255 260 265 270	818
CGT AAA GGC TTT GCT CAG GTT GCA ATC GAT GCA AAA GTG CCC ATT ATT Arg Lys Gly Phe Ala Gln Val Ala Ile Asp Ala Lys Val Pro Ile Ile 275 280 285	866
CCT ATG TTT ACA CAA AAC ATC CGA GAA GGA TTT AGA TCA CTC GGA GGA Pro Met Phe Thr Gln Asn Ile Arg Glu Gly Phe Arg Ser Leu Gly Gly 290 295 300	914
ACA AGA TTG TTT AAA TGG CTT TAT GAA AAA TTC CGC TAT CCG TTT GCT Thr Arg Leu Phe Lys Trp Leu Tyr Glu Lys Phe Arg Tyr Pro Phe Ala 305 310 315	962
CCA ATG TAT GGA GGT TTT CCT GTG AAG TTG CGG ACC TTC TTG GGT GAT Pro Met Tyr Gly Gly Phe Pro Val Lys Leu Arg Thr Phe Leu Gly Asp 1010	

320	325	330	
CCT ATT CCG TAT GAC CCA GAG GTA ACA GCA GAA GAA TTA GCT GAA AAG Pro Ile Pro Tyr Asp Pro Glu Val Thr Ala Glu Glu Leu Ala Glu Lys 335 340 345 350			1058
ACT AAG AAC GCT GTT CAA GCT TTG ATC GAC AAG CAC CAA AGG ATA CCG Thr Lys Asn Ala Val Gln Ala Leu Ile Asp Lys His Gln Arg Ile Pro 355 360 365			1106
GGG AAC ATT AGG AGT GCT TTG CTG GAC CGC TTT CAT AAA GAG CAG AAG Gly Asn Ile Arg Ser Ala Leu Leu Asp Arg Phe His Lys Glu Gln Lys 370 375 380			1154
GCT AAT TGACAGATAA TTTCGTGTGT CTCTAATGCC AAGTTTGCAT CTGTGGTACT Ala Asn			1210
GCCTTTTAAA CTTTGTAGG TTGTATAATA GTCTTTTAA AACCTGCTAA TGAGCATCTT			1270
TCTTTAGAAT TTAACTTTC TATGGTTAAT CTTACCCTTT CAATCAGTTT TGTCAAATGT			1330
AGTAAGTAAC TCACCATTTC CTGACTCAG AAGATGATCA TGTTCATTT GCAATCCTTA			1390
GTAGTATATG ATAGGCACAA GTCTTCCTGG ATAACAGCAG TCTAGGTTAG GCAGACTTCA			1450
CTAAATTATG CTTCTGTGTT TGTGTCTGTT GTGCCAATGG GATGGTGAGA GGTACCTGAG			1510
CTGGGGTGTA GCACTTGCTT AGTTTAGGTG AGCGGTGCCT TTGTATCCAT CCTTCATTCC			1570
ACTCACAAC GCTGGTAAGT TGTGGAACGC TTTTGCACCT CCGTGGTTTA TGTTAAGCAT			1630
GGTAATACTG TGGTGGTGAA CATTAAGTTG CTTATGTCTA CTGTGATATA TCCGTATCTA			1690
TATATCTGAT AGATATCTGT AGTGAGAAAT CTGTAATTGA ATTGATTCCT GAATCCTTTA			1750
GGAGTTATGT ACATTTTGT CCATTGGGTA TTGAAATTAC TCCTTCAGGG ATGATGTCCA			1810
GTTGCCAAGC ACAATGAAAG AGATGTACTA TTTTGTAAT CAAATCATTG TTACTTGTAG			1870
TATTTATTAA CAGAGTTGTT TGGTATTTAT TTAAACAAG GATCTAAGTA AATGTATAGG			1930
ATGGATTGTC AGCACGTGTC TTACACGGTG TAAAGATAAA ATCGTGGGCA CGCTCTGCAG			1990
TTTATGTTAG AGCAGATCCT AGGCTGAGGA GTGACAACTG AGGAAGCTCT GCCTCGCTTA			2050
TCCCAGGGCT CGCTTCCGTG CTCTGCAGCC AGTCACGTGG TTATATCTGT AAAATGTACT			2110
TGAATGATGA ATTTCTGTAA ACATGGAGTC CTCTGAAAAT GTAATTACTT TGAATTGTTG			2170
TCTTCTTGAA AACTTATTTT AAAAGTATTT AAACAGCCAA GTATCATGAA TTCATCGCAG			2230
TGTGTCTTCC CCCTTTCAG TACTTGGGTT TTATATGGAC TTTTAACTA CAAGTGTATA			2290
TTTGTGATAT GTATAATGTA TGTATAATGT ATATAGAAAT GGTGTGAAAA TAAAGCTGAA			2350
CTGCAACTCC TCTCTAGTAA AAA			2373

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 384 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```

Asp Phe Leu Val Arg Phe Thr Pro Ala Ala Ala Gly Leu Cys Ser Leu
 1           5           10           15
Thr Glu Trp Cys Cys Val Thr Trp Ile Cys His Phe Tyr Ser Val Ile
          20           25           30
Ser Trp His Cys Ser Asn Thr Tyr Gly Asn Glu Glu Arg Asp Arg Ile
          35           40           45
Trp Asp Phe His Leu Arg Lys Met Ile Asn Asn Asn Gln Thr Cys Ala
          50           55           60
Ala Gly Gln Asp Ser Val Pro Tyr Val Thr Cys Met Leu His Val Leu
          65           70           75           80
Glu Glu Trp Leu Gly Val Glu Gln Leu Glu Asp Tyr Leu Asn Phe Ala
          85           90           95
Asn His Leu Leu Trp Val Phe Thr Pro Leu Ile Leu Leu Ile Leu Pro
          100          105          110
Tyr Phe Thr Ile Phe Leu Leu Tyr Leu Thr Ile Ile Phe Leu His Ile
          115          120          125
Tyr Lys Arg Lys Asn Val Leu Lys Glu Ala Tyr Ser His Asn Leu Trp
          130          135          140
Asp Gly Ala Arg Lys Thr Val Ala Thr Leu Trp Asp Gly His Ala Ala
          145          150          155          160
Val Trp His Gly Tyr Glu Val His Gly Met Glu Lys Ile Pro Glu Gly
          165          170          175
Pro Ala Leu Ile Ile Phe Tyr His Gly Ala Ile Pro Ile Asp Phe Tyr
          180          185          190
Tyr Phe Met Ala Lys Ile Phe Ile His Lys Gly Arg Thr Cys Arg Val
          195          200          205
Val Ala Asp His Phe Val Phe Lys Ile Pro Gly Phe Ser Leu Leu Leu
          210          215          220
Asp Val Phe Cys Ala Leu His Gly Pro Arg Glu Lys Cys Val Glu Ile
          225          230          235          240
Leu Arg Ser Gly His Leu Leu Ala Ile Ser Pro Gly Gly Val Arg Glu
          245          250          255
Ala Leu Leu Ser Asp Glu Thr Tyr Asn Ile Ile Trp Gly Asn Arg Lys
          260          265          270
Gly Phe Ala Gln Val Ala Ile Asp Ala Lys Val Pro Ile Ile Pro Met
          275          280          285
Phe Thr Gln Asn Ile Arg Glu Gly Phe Arg Ser Leu Gly Gly Thr Arg
          290          295          300

```

Leu Phe Lys Trp Leu Tyr Glu Lys Phe Arg Tyr Pro Phe Ala Pro Met  
 305 310 315 320  
 Tyr Gly Gly Phe Pro Val Lys Leu Arg Thr Phe Leu Gly Asp Pro Ile  
 325 330 335  
 Pro Tyr Asp Pro Glu Val Thr Ala Glu Leu Ala Glu Lys Thr Lys  
 340 345 350  
 Asn Ala Val Gln Ala Leu Ile Asp Lys His Gln Arg Ile Pro Gly Asn  
 355 360 365  
 Ile Arg Ser Ala Leu Leu Asp Arg Phe His Lys Glu Gln Lys Ala Asn  
 370 375 380

## (2) INFORMATION FOR SEQ ID NO:39:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2241 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

(B) CLONE: HW047

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 8..511

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION: complement (1675..1833)  
 (D) OTHER INFORMATION: /label= SAC\_23915

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTCGACC AGG GCC TTC CAT CAG ACC TGC CCA AAC TTT GTC ATA GAG AAG Arg Ala Phe His Gln Thr Cys Pro Asn Phe Val Ile Glu Lys 1                  5                  10	49
ATT GAA AGG ATT CAA AAT CCA GCT CTT TGG AGG AGG TAC CAG GCG TAC Ile Glu Arg Ile Gln Asn Pro Ala Leu Trp Arg Arg Tyr Gln Ala Tyr 15                  20                  25                  30	97
AAA AAA ATT ATG GAT GAA AAG AAT GGC AAT GTG ATA AAT GAG AAG CAA Lys Lys Ile Met Asp Glu Lys Asn Gly Asn Val Ile Asn Glu Lys Gln 35                  40                  45	145
CTC TTC CAT GGG ACG GAG TTT GGC TCT CTG GCA CAG CTC AAC AGT AAT Leu Phe His Gly Thr Glu Phe Gly Ser Leu Ala Gln Leu Asn Ser Asn 50                  55                  60	193
GGA TTT AAC CGC AGC TAT GCT GGG AAG AAC GCT ACG GCT TAT GGG AAA Gly Phe Asn Arg Ser Tyr Ala Gly Lys Asn Ala Thr Ala Tyr Gly Lys	241

65	70	75	
GGA ACC TAT TTT GCA GTC AAT GCT TCC TAT TCT GCC CAT GAC ACA TAT			289
Gly Thr Tyr Phe Ala Val Asn Ala Ser Tyr Ser Ala His Asp Thr Tyr			
80	85	90	
TCC AAA CCA GAT GCC AAT GGG AAA AAG TAT ATG TAT TAT GTG CGG GTC			337
Ser Lys Pro Asp Ala Asn Gly Lys Lys Tyr Met Tyr Tyr Val Arg Val			
95	100	105	110
CTC ACT GGA AAC TAC ACA CAA GGA AAC GCA TCA CTG ATT GTG CCT CCT			385
Leu Thr Gly Asn Tyr Thr Gln Gly Asn Ala Ser Leu Ile Val Pro Pro			
115	120	125	
TCA AGG GAC CCT CAA AAT GCT ACT GAC TTG TTT GAC ACT GTC GCA GAT			433
Ser Arg Asp Pro Gln Asn Ala Thr Asp Leu Phe Asp Thr Val Ala Asp			
130	135	140	
AAT GTT ATA AAT CCG AGT ATA TTT GTA GTG TTT TAT GAC AAT CAA GCA			481
Asn Val Ile Asn Pro Ser Ile Phe Val Val Phe Tyr Asp Asn Gln Ala			
145	150	155	
TAC CCT GAG TAT CTT ATC ACA TTT AGG AAG TAACACTTTG GGGTGCCTTC			531
Tyr Pro Glu Tyr Leu Ile Thr Phe Arg Lys			
160	165		
TGAAAGACAT TGGTATCTGC CGTAGGACAA ATTTTAGCTG TTTTCCTTCT TAGACTTTCT			591
TCTCTAAGAG CACGAGGTCT TTCTCTGCCA CGGATTGTTT CCTCAGCTGC CCCTTCATAG			651
TGAACATGAA CCTACTTGAG GGCCAACCAC TTGAGAATTT CAAGCAGGTG ACAGCTTTAG			711
CTGCCTTCAC AAATGTGAAC CTCAAATCCA CGGAAAAGAA CAAACTTACG TTCTGGAAGG			771
ACAATAAAAC CTTTATGCAC TTTTCAGAAAG CCCACAATAC ACCCACCACC AGTGGAGTTA			831
CATCAGAGCA CCACTGTGCT GAGGGGCACC CAGGACCACT GAGCCAGTAA GGCTGCCAAC			891
TGAACAGATT CCAAGACAGA GAACTTCGTT CATCACTAAC AGCGCCGTGT CCTGTGGGTT			951
GGACAGTTCC AGCCAGGGTT ATGTACAGAC TGGGTCTCTGG GATGTGGCTC TGGGTCTGTG			1011
TGAACTGAGC CCCCCAAGGA GAGGCACAGC TGAAGCCCTG TCAGGGGGCC CTTTTAATGT			1071
TGACACCAAT GTTGAGGCTA CCTAGTGAAG AAATGGGATG GGGCCTGCTA CCTCTCCTGT			1131
GGAAGTCCTT CTACTGGGTG AAAATGAATG AACTCTAACA GTATTGGCTA GCGTCTTTCT			1191
TGTGGCAGGC TGGCTTGTGA TAACTTGAAG CTTTATGACT AAGTCAGTCA TTTTGCTAAC			1251
TCACTAACAG TATCTACTGG ATGCGATTTG TTTCTTGTA TGCCAGAGAT CACCCAGAGC			1311
CTCACATGGG CCAGGCAAGC ACTCTACTAC TGAACTATAT TCCCAGCCTC TTAGTGAGGC			1371
AAGACTCCAC TAAGTTGCCC TGGCAGGTTT GAACTCATTA TCTTCCTGCA TTGGCTACTT			1431
ATTGGGGTGG GGTGGGGTGC TTAAATTGTT CTTTTAATTT TCTCCTGTAC TCCCATTTAT			1491
TGAGTGCCCTG GACCCTTGTC AGGTAGAATA AGAGGTGTGT TCACATAGAT AATTTTGT TT			1551
CCATAAAAAA AAATGCACTG TAAGAATTAT GTCAAGGCCT AGGAGCAGGC TCAGAGTACA			1611
AGAGCTGTCC TTGCCAAAGG TCACTAAGGC AGGAACCGCA GAGCCAGCCA CCAGAGTCTT			1671



GGTACTGGGT TGAGCCCCTC ACTATTCCCA GCACAGGGCC TGGCCCTGAG GTAAGCCTGT 1731  
 GCTGCTTCAG AAGTTGGCTT TTCTCCTTCT GGGATTGGTG TCCACTCTGC CCCTTTTCTA 1791  
 GGTATGGAG CTAGTGGGT GTAATGTTCC CAGGAGGTAG GTACACTCTT AGGTGTTGTG 1851  
 GGGATCCAAG TGCTGGTTC CACCATTGGC ATTAACAGTG GACTGCAAGC AGAGAGAGGG 1911  
 GTCCCCATCA TTGTTTCGTG GACACTGAGG CCTTGAGACC TTGGCTGCTG TGTCTATGTG 1971  
 GACTGGCAAA GACTGTGGCT TGGGAGGGGC TTCCTCAGCA CTGGCACCTG AACTCAGCAC 2031  
 CAAGGATGGC CCCAAGAGGC TTGTTTTCAT TTGCCTTTGT TCTCCAGTGC TTTGGTTACC 2091  
 TCAGGGCAGG GCTGAACAAT AGTGTGATGT TACATTGAAG AATAAGACC CAACCAACAA 2151  
 ATGTTAGTGT CATTTTCAGT CTAAGTAAAT AAGCCATTAT GTCATTCCC TTACTIONCTA 2211  
 GTAAAGTTAT GTAATTACAT AATGTAAAAA 2241

## (2) INFORMATION FOR SEQ ID NO:40:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 168 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Arg Ala Phe His Gln Thr Cys Pro Asn Phe Val Ile Glu Lys Ile Glu  
 1 5 10 15  
 Arg Ile Gln Asn Pro Ala Leu Trp Arg Arg Tyr Gln Ala Tyr Lys Lys  
 20 25 30  
 Ile Met Asp Glu Lys Asn Gly Asn Val Ile Asn Glu Lys Gln Leu Phe  
 35 40 45  
 His Gly Thr Glu Phe Gly Ser Leu Ala Gln Leu Asn Ser Asn Gly Phe  
 50 55 60  
 Asn Arg Ser Tyr Ala Gly Lys Asn Ala Thr Ala Tyr Gly Lys Gly Thr  
 65 70 75 80  
 Tyr Phe Ala Val Asn Ala Ser Tyr Ser Ala His Asp Thr Tyr Ser Lys  
 85 90 95  
 Pro Asp Ala Asn Gly Lys Lys Tyr Met Tyr Tyr Val Arg Val Leu Thr  
 100 105 110  
 Gly Asn Tyr Thr Gln Gly Asn Ala Ser Leu Ile Val Pro Pro Ser Arg  
 115 120 125  
 Asp Pro Gln Asn Ala Thr Asp Leu Phe Asp Thr Val Ala Asp Asn Val  
 130 135 140  
 Ile Asn Pro Ser Ile Phe Val Val Phe Tyr Asp Asn Gln Ala Tyr Pro  
 145 150 155 160

Glu Tyr Leu Ile Thr Phe Arg Lys  
165

## (2) INFORMATION FOR SEQ ID NO:41:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1918 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW049

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1289..1612)
- (D) OTHER INFORMATION: /label= SAC\_24337

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GTCGACGTGG TGATTATCTT AAAGGTAAGG GAGGTTTTTT TTTCCCTGAT ACTGATGAAG	60
GCTGAGTGTG ACTGTTTAGT ATAGCTAGTG GGCTTTGTCA CTTAAGCCTT CCTTAGCATG	120
TTCCCTTTCC CTTATTTTTC CTGGATACTT ACGTGCTAAC GTGCTAACAA CTCCAGATCA	180
AACTGAGTAT GGAGATTTTT GCTTTAACAA TTTAGGAGTG GCAAACAATT TATAACTGGC	240
AGGGTTTGAG GACAGCCGTG TTGAAAAACC AAAAGAAAAT GCTAAGGCAG GTCTTGACCA	300
ATGTGGATAC TTTGACTAAT TAACTTCGTT ACTTTAGGAG GAGGTTGTTC CTGCACGCCA	360
TCCATCTCCA GCGTTCGAAG GTTAGATTGT TGGGGTGGTG TGTCTGTTC ACCAGATTGG	420
CTGTTTCTGT GATGAATCAA CTAGCTCACT ATGACTGCTA ATGAAAACAC TGGAACACCT	480
GAGAACTGA AGGGTGGCGT GTTGGGGCAA GGGTGCTTAC ATATCAAGCC TTTTCAATAG	540
TTTGTTTTTC TCCAGGTTCT GTTAGCAGTG CTGTTACTTG GTGCTTCAGA GGTAAGTAAT	600
TGCAGCCAGC GTGCAGTGGT GTGCACACTG CTGGACTCAG GTCCTAAGTA CTTCATATAT	660
ATGATGACTT TCATAGAATC TCGTTCGGCT GATGATTGCT GTTGAGACTT GGAAATCTGA	720
TTTTTCTGAG GATGATGGGT TGGATGGGAA CAATCTGCCA GGATTTTTAA CATTGCCCCAT	780
CTCGTGTGTT TCAGATATGG TGA CTGGCCG AGCGCAAAGG AGCCATGCCA GGTAAATGTT	840
CTAGCAATAT TATTGTTTAT TAGGTAGATA CGTAATGGCT TGATCCAGTG CTTTCCTGCT	900
CACCCGTATG AACTGAATAC CGCCCCAGTC TGATAGCTGT GGAGAAAGGT ATTTTCTGAG	960
TTGAGCTCAT CTTGGATCAA ACAGGAAGGG TTTTATGCTA GTTAGTATAT GATCTCTGTA	1020

GGATACGTTT AGGCAATAAG GATAATGACC CAAATTTATT TACTGTTTTG CAGGAAAAC 1080  
 AAGAGCACAT CTTCATAAAG TCCAGGCTAA AGTAAGGTGA GAAACTAAG ATGTAAAGTG 1140  
 TACCAAGTTA GTAACCTCCC TATGAAGAAA CATTACACAG TCTTACTTCC TGTCCCTAGCT 1200  
 CCAGAGCCTG AAAAGGTGAA CCCACTGGGG CTGGCTGGGG GAAAAGAGGA AATTTTGTTC 1260  
 CAGAAGGAAC TGTCTGAGGG ATCATTGTGA TGTGTGGAAA GGGAAGAATT AAACTTTGAT 1320  
 TAAAAAGGAA ACAAGGTGGT TTGTCTGTTT TTCAGGTCTG AGGAGAGGGA TGAGGTGTGT 1380  
 GTTAGAGGTA GGTAGCCCTT CAAGTCATTG TAATTAGGTG GGCAGGGTAG ATTTTGCTTT 1440  
 CTGCTAATTA AAACTTTAC CTTACAGGAA CTGAAATCAA AGATCCGATT ACCTAGGCCA 1500  
 GTGGCACGGC AAACCTCAGA ACTCATGGAA TACAGCCTGT TGATTCCTTG TGGATTTTAA 1560  
 GCATCTGAGC AGTATCAGTG GCACTCCAAG TGAAATGGTT GATCTCTTCA GTACCGCTGG 1620  
 AGCAGATTAA TATTTATTTT ACTGCCAGAG GCTATATTTT TCTGCCCCC CATACCCCT 1680  
 TCCTCTTCTC CTTAGACTAC TTAAGATTGA AATATTTGCA ACTCCAAATT TATCTTCCTT 1740  
 GCTTACCACT TTTTTTCCG AGCTGGGGAC CAAACCCAGG GCCTTGCGCT TCCTAGGTAA 1800  
 GCGCTCTATC ACTGAGCTAA ATCCCCAGCC CCGCTTACCA CATTGTTAAT CTGATCCTTC 1860  
 CGCAGTTATT TGTGCTGTTC TTTTGTTC TGGCAATAAA TGCTTTTAT GCTAAAAA 1918

## (2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3942 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

- (B) CLONE: HW050

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..1326

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (3542..3733)
- (D) OTHER INFORMATION: /label= SAC\_24644

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GTCGAC ATA AAT ATG GCC GTC CGG GTA CCC CTC GGC TGC ACG GGC TCC 48  
 Ile Asn Met Ala Val Arg Val Pro Leu Gly Cys Thr Gly Ser  
 1 5 10  
 TTC TGC CCG CGG CTG CTA CCA CTG CTG GCA TTG TTG GAG CTG CTC GTC 96

Phe 15	Cys	Pro	Arg	Leu	Leu 20	Pro	Leu	Leu	Ala	Leu 25	Leu	Glu	Leu	Leu	Val 30	
GAC Asp	CCC Pro	AGC Ser	TTG Leu	GGC Gly 35	CGC Arg	GTC Val	CAT His	CAC His	CTG Leu 40	GCG Ala	CTC Leu	AAA Lys	GAT Asp	GAC Asp 45	GTG Val	144
CGG Arg	CAC His	AAA Lys 50	GTC Val 50	CAT His	CTC Leu	AAC Asn	ACC Thr 55	TTC Phe 55	GGC Gly	TTC Phe	TTC Phe	AAG Lys 60	GAC Asp 60	GGG Gly	TAC Tyr	192
ATG Met	GTG Val	GTC Asn 65	AAC Ile	ATC Ser	AGC Ser	AGC Leu 70	CTC Ser	TCT Ser	GTG Val	AAC Asn	GAG Glu 75	CCT Pro	GAG Glu	GGA Gly	GTC Val	240
AAG Lys	GAC Asp 80	AAA Lys	GAC Asp	ACT Thr	GAG Glu	ATT Ile 85	GGC Gly	TTC Phe	AGT Ser	CTT Leu	GAT Asp 90	CGG Arg	ACC Thr	AAG Lys	AAT Asn	288
GAT Asp 95	GGC Gly	TTT Phe	TCT Ser	TCT Ser	TAC Tyr 100	CTG Leu	GAT Asp	GAA Glu	GAT Asp 105	GTG Val	AAT Asn	TAC Tyr	TGT Cys	ATT Ile	TTA Leu 110	336
AAA Lys	AAA Lys	AAG Lys	TCC Ser	ATG Met 115	TCT Ser	TCT Ser	GTC Val	ACT Thr	CTG Leu 120	CTC Leu	ATC Ile	TTA Leu	GAC Asp	ATC Ile 125	TCC Ser	384
GGA Gly	AGT Ser	GGA Gly	GTC Val	AAG Lys 130	GTC Val	AGA Arg	TCC Ser	CCA Pro	CCA Pro	GAA Glu	GCT Ala	GGC Gly	AAG Lys 140	CAG Gln	TTA Leu	432
CCC Pro	GAG Glu	ATT Ile 145	GTC Val	TTC Phe	AGC Ser	AAG Lys	GAT Asp 150	GAG Glu	AAA Lys	GTC Val	CCG Pro	AGT Ser 155	CGG Arg	AGC Ser	CAG Gln	480
GAG Glu 160	CCT Pro	GCT Ala	GTC Val	AGC Ser	TCC Ser	AAC Asn 165	CCC Pro	AAA Lys	GAC Asp	AGC Ser	AAG Lys 170	GTG Val	CAG Gln	AGA Arg	ACC Thr	528
CCT Pro 175	GAT Asp	GGT Gly	TCC Ser	AAG Lys	GCT Ala 180	CAA Gln	CGA Arg	AGT Ser	ACA Thr	GTG Val 185	GAT Asp	TCA Ser	AAG Lys	ACT Thr	ATA Ile 190	576
GCA Ala	GAG Glu	AAA Lys	TTC Phe	TTC Phe 195	TCA Ser	ATT Ile	CAT His	AAG Lys	AAT Asn 200	GAT Asp	GGG Gly	GCA Ala	GTT Val	TCA Ser	TTT Phe 205	624
CAG Gln	TTC Phe	TTC Phe	TTT Asn 210	AAC Ile	ATC Ser	AGC Thr	ACC Ser	AGT Lys 215	GAC Asp	CAG Gln	GAA Glu	GGC Gly	CTC Leu 220	TAC Tyr	AGC Ser	672
CTT Leu	TAC Tyr	TTC Phe 225	CAC His	AAG Lys	TGC Cys	CCA Pro	AGC Ser 230	AGC Ser	AAG Lys	CTG Leu	AGG Arg	TCT Ser	GGT Gly	GAG Glu	CAG Gln	720
GTC Val	TCA Ser	TTC Phe 240	AGC Ser	CTA Leu	AAT Asn	ATT Ile 245	GAT Asp	ATC Ile	ACC Thr	GAG Glu	AAG Lys 250	AAT Asn	CCT Pro	GAC Asp	AGC Ser	768
TAC Tyr 255	CTG Leu	TCT Ser	GCA Ala	GGG Gly	GAA Glu 260	ATC Ile	CCT Pro	CTC Leu	CCC Pro	AAG Lys 265	TTA Leu	TAC Tyr	GTT Val	TCC Ser	ATG Met 270	816

GCC TTG TTG TTC TTC CTG TCT GGG ACC GTG TGG ATT CAC ATC CTT CGT Ala Leu Leu Phe Phe Leu Ser Gly Thr Val Trp Ile His Ile Leu Arg 275 280 285	864
AAA CGA AGG AAT GAT GTA TTT AAA ATT CAC TGG TTG ATG GCG GCC CTT Lys Arg Arg Asn Asp Val Phe Lys Ile His Trp Leu Met Ala Ala Leu 290 295 300	912
CCT TTC ACC AAG TCT CTC TCC TTG GTG TTC CAT GCA ATC GAC TAC CAC Pro Phe Thr Lys Ser Leu Ser Phe Val Phe His Ala Ile Asp Tyr His 305 310 315	960
TAC ATC TCC TCG CAG GGC TTT CCG ATT GAA GGC TGG GCT GTT GTG TAC Tyr Ile Ser Ser Gln Gly Phe Pro Ile Glu Gly Trp Ala Val Val Tyr 320 325 330	1008
TAC ATA ACT CAC CTG CTG AAG GGT GCA CTG CTG TTC ATC ACC ATC GCC Tyr Ile Thr His Leu Leu Lys Gly Ala Leu Leu Phe Ile Thr Ile Ala 335 340 345 350	1056
CTC ATT GGC ACT GGC TGG GCC TTC ATC AAG CAC ATC CTG TCT GAT AAG Leu Ile Gly Thr Gly Trp Ala Phe Ile Lys His Ile Leu Ser Asp Lys 355 360 365	1104
GAC AAG AAG ATC TTC ATG ATT GTC ATT CCG CTC CAG GTA CTG GCG AAT Asp Lys Lys Ile Phe Met Ile Val Ile Pro Leu Gln Val Leu Ala Asn 370 375 380	1152
GTG GCC TAC ATC ATC ATA GAG TCT ACT GAG GAG GGC ACA ACT GAG TAT Val Ala Tyr Ile Ile Ile Glu Ser Thr Glu Glu Gly Thr Thr Glu Tyr 385 390 395	1200
GGC TTG TGG AAG GAC TCT CTA TTC CTG GTG GAT TTG CTG TGC TGC GGC Gly Leu Trp Lys Asp Ser Leu Phe Leu Val Asp Leu Leu Cys Cys Gly 400 405 410	1248
GCC ATC CTC TTC CCA GTG GTG TGG TCA ATC AGG CAT TTA CAA GAA GCC Ala Ile Leu Phe Pro Val Val Trp Ser Ile Arg His Leu Gln Glu Ala 415 420 425 430	1296
TCG GCC ACA GAT GGA AAG CTG CCA TTA ACC TAGCAAAGCT GAAGCTTTTC Ser Ala Thr Asp Gly Lys Leu Pro Leu Thr 435 440	1346
AGACATTACT ACGTCCTGAT CGTGTGCTAC ATCTACTTCA CCAGGATCAT CGCCTTTCTT	1406
CTCAAGTTTCG CTGTTCCCTT CCAGTGGAAG TGCTCTACC AGCTGCTGGA TGAAACAGCC	1466
ACACTGGTGT TCTTTGTCTT GACGGGGTAT AAATTTGCGC CAGCTTCAGA TAACCCCTAC	1526
CTGCAGCTCT CTCAGGAAGA AGATGACCTG GAGATGGAGT CTGTAGTGAC GACGTCAGGA	1586
GTAATGGAGA ACATGAAGAA GGTGAAGAAA GTGAGCAATG GTGCTGTGGA GCCCCAGGGC	1646
AGCTGGGAAG GCACCGCGTG ACAGCGCAGC TGAGGACCGC AGGCCGTCAA TGAACCTAGC	1706
TTATCCATAG CCCTATCAGT GAGCACGCCT GTTCTTGACA GACGATGGCA TCTCCCAGCA	1766
GCGACACCCC GCGCTCAGC ACCATGGCCG AGCCGAGTGC TATGGACACA CACTTTTGTA	1826
CTCTTAGGGA TCTGGATTGA GGTGGGCTAC AGGCACTGAG GACTGCCCTC AGTGAGGGCG	1886
GAGGCTTAGA GGAGGAGTAA AGACTGTCTG TCTGATTAC CTTGCTAATT GGGAGCCTGG	1946

GTCTGTGCAG	GGGAGCCCCC	AGTGGCTGCT	ACTCAGAAAG	GCTGTGGGTG	TGGCACGGGT	2006
CTGCCACTCG	CTGGTCACAG	TTCTGTGGAG	GCTGGGAGGA	TGTGACTCAC	TGCTTGCCCT	2066
AGCATTTACG	AGGCTTTGGG	TTCAATTCCT	AGTGCCAGGG	TGGGGGGAAA	AGAAGTTACT	2126
TTGGGACTAG	TCAGGTATCT	ACCAAGGGCA	CAGAGAAGAG	AAAGAGATTG	TCAGCCTTTG	2186
GGAACAAGAA	CAACAGAAAT	GATTGAAATT	CTTTTAAATA	GCCAATAAAC	CCCTAGGTGG	2246
CAGAAGACAC	CTGTCCAGGG	TCACTGGCCC	CCAGGCACCA	CTCTAAATGT	CCCTGCCCCCT	2306
GATGTGACCT	ACAACAGTGG	TGCTAGAGGC	CAGAGGGAGG	GAGGGAGGGG	GCCTGTTACA	2366
TGGCTGGCTG	CAGGGGAGAA	GAGCCTTGGC	TCTCCAGAAG	TGTTCTCCAG	TTGGCTCTCA	2426
GGACTCCATC	CCCTGGCTGC	CTGCCCAGAG	CAGCCGCCTG	GCTTCCTCTC	CTTCCACCTG	2486
TGTCTTCTGA	TCCGAGAAAA	CAGTTTTTCCT	GAAGGAGCAG	CCGAGGACAG	AAAGGGTAGT	2546
CTCTAAACAA	AACTTAGGT	TAAGCATCTT	CCACAGCTCG	CATGGGCCAG	GCAGGACCTC	2606
TGAGCCTCAG	CTATAGTCCA	AGCATGCGTG	AAGGAAGTGT	TTGAGCCACT	CATGCTCAGA	2666
GACCTGAGGC	AGCCGGGTCT	GCTCAGAGCT	TGGACAGGGA	CACCAGCGGC	ACCTGCGGAA	2726
GGCCTGTGAG	GGGCCACAGG	AGACCAGCTC	AGCACCTCAG	AGCTGCTGCC	TGCCCAGAAC	2786
TGGCGGGCGA	GTTGCACCTT	CAAGGAACAG	GCACTTCCGG	GAGGGGGGTG	AGTTACTCCA	2846
CCTCCCAGTG	CAGAAGACAA	AGGTCTGGTC	TCCACACCAC	ACCCCAAGTC	CCTGTCTCTGT	2906
GCCACCAGTG	AGAAGCTGCT	GTCTGCTGCC	CTAGCAGGTC	CGTCACTACC	GCAGGTGCTC	2966
CATCTTGAGA	AAGGTCGACA	GGCCTGAAGG	ATCCACCTAA	AGGACCTTTT	TCTCCTGTTT	3026
TGTTTGCTTT	GTTTATACTT	GTTTAAAGTA	TGTGGTGGGA	AGGGAGTGTC	GAGGAGACAA	3086
GGTGTTAATG	CTTTGTCTTC	TGCACATAAG	CACTCCTCCA	GCACTGCAGA	AGTGTGTCCC	3146
TGTGGTGA CT	CTAAGCCACA	GCTGTCCCTC	AGAGGGCGCG	ATCTTTATGG	AAGGAGCCTC	3206
CTTTCCTTGG	CTGGGTTTTA	ATAAGGCCAC	TCTGACAGAG	AGCAGAGGGA	GCTGTGTACA	3266
ACCCAGCCAC	AGAGCTCTCT	GGGTGCACAG	CTGAGCGGCT	ATGGGAGACA	GCCCGCTCCT	3326
TCCAGAGCAT	GGGTGCACAG	GCAGGGAGGT	GTGCTACAGC	AGCGGCCCTC	CCAAGTCCTG	3386
CCCTGTGACT	GCTGTAGGTG	GGAAGCCAGG	CATGTTGGGG	GTTGTGGTGG	GGAGGGGGGA	3446
TCAGAAGCCT	AACATGGGGC	TGGGAAGAAC	TGGCCTGGTG	CCCAGGCTGT	GGAAAGGAAA	3506
ACTCCCAGCC	CTCAGCATGC	ACTGTAAGGT	GTCTGTCTCA	TCTGCATCGC	TTGAGAGCTG	3566
GGAGCCC GCA	TTCAGCAGTG	TCCTTGGGTA	GAGGAGTCTG	ATTGGCATCC	AGCCGCCCCA	3626
GTTTCATGAGC	TGCCCTCTGC	AGCCTTGAGT	CTTAGCCTAC	CCAGCTCCAC	AACCGTGTTT	3686
GAGTGGTGAA	GCCCTTG TGT	TGGAGGGGCT	GAGCAGAAAT	GGGGTG TACA	TTGCACTGCC	3746
ACTGTGGTAT	TAATAGCTCT	GTGTTGTGTC	CCACCCCCCA	CCCCCATCT	TCTGCCTGCG	3806

TGAGGTATCC ATACTTGT TTGAGAGTGCT GTGGACTGGA AGCTGTGGGC TGTGACATTT 3866  
 AATAAAGTGT GGATTTTGTT AGATAATTAT TTCTTGACA ACAGGAACAA GTGTCAGCT 3926  
 CTGAAAAAAA GTCGAC 3942

## (2) INFORMATION FOR SEQ ID NO:43:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 440 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Ile Asn Met Ala Val Arg Val Pro Leu Gly Cys Thr Gly Ser Phe Cys  
 1 5 10 15  
 Pro Arg Leu Leu Pro Leu Leu Ala Leu Leu Glu Leu Leu Val Asp Pro  
 20 25 30  
 Ser Leu Gly Arg Val His His Leu Ala Leu Lys Asp Asp Val Arg His  
 35 40 45  
 Lys Val His Leu Asn Thr Phe Gly Phe Phe Lys Asp Gly Tyr Met Val  
 50 55 60  
 Val Asn Ile Ser Ser Leu Ser Val Asn Glu Pro Glu Gly Val Lys Asp  
 65 70 75 80  
 Lys Asp Thr Glu Ile Gly Phe Ser Leu Asp Arg Thr Lys Asn Asp Gly  
 85 90 95  
 Phe Ser Ser Tyr Leu Asp Glu Asp Val Asn Tyr Cys Ile Leu Lys Lys  
 100 105 110  
 Lys Ser Met Ser Ser Val Thr Leu Leu Ile Leu Asp Ile Ser Gly Ser  
 115 120 125  
 Gly Val Lys Val Arg Ser Pro Pro Glu Ala Gly Lys Gln Leu Pro Glu  
 130 135 140  
 Ile Val Phe Ser Lys Asp Glu Lys Val Pro Ser Arg Ser Gln Glu Pro  
 145 150 155 160  
 Ala Val Ser Ser Asn Pro Lys Asp Ser Lys Val Gln Arg Thr Pro Asp  
 165 170 175  
 Gly Ser Lys Ala Gln Arg Ser Thr Val Asp Ser Lys Thr Ile Ala Glu  
 180 185 190  
 Lys Phe Phe Ser Ile His Lys Asn Asp Gly Ala Val Ser Phe Gln Phe  
 195 200 205  
 Phe Phe Asn Ile Ser Thr Ser Asp Gln Glu Gly Leu Tyr Ser Leu Tyr  
 210 215 220  
 Phe His Lys Cys Pro Ser Ser Lys Leu Arg Ser Gly Glu Gln Val Ser  
 225 230 235 240

Phe Ser Leu Asn Ile Asp Ile Thr Glu Lys Asn Pro Asp Ser Tyr Leu  
                   245                                  250                                  255  
 Ser Ala Gly Glu Ile Pro Leu Pro Lys Leu Tyr Val Ser Met Ala Leu  
                   260                                  265                                  270  
 Leu Phe Phe Leu Ser Gly Thr Val Trp Ile His Ile Leu Arg Lys Arg  
                   275                                  280                                  285  
 Arg Asn Asp Val Phe Lys Ile His Trp Leu Met Ala Ala Leu Pro Phe  
                   290                                  295                                  300  
 Thr Lys Ser Leu Ser Leu Val Phe His Ala Ile Asp Tyr His Tyr Ile  
 305                                  310                                  315                                  320  
 Ser Ser Gln Gly Phe Pro Ile Glu Gly Trp Ala Val Val Tyr Tyr Ile  
                   325                                  330                                  335  
 Thr His Leu Leu Lys Gly Ala Leu Leu Phe Ile Thr Ile Ala Leu Ile  
                   340                                  345                                  350  
 Gly Thr Gly Trp Ala Phe Ile Lys His Ile Leu Ser Asp Lys Asp Lys  
                   355                                  360                                  365  
 Lys Ile Phe Met Ile Val Ile Pro Leu Gln Val Leu Ala Asn Val Ala  
                   370                                  375                                  380  
 Tyr Ile Ile Ile Glu Ser Thr Glu Glu Gly Thr Thr Glu Tyr Gly Leu  
 385                                  390                                  395                                  400  
 Trp Lys Asp Ser Leu Phe Leu Val Asp Leu Leu Cys Cys Gly Ala Ile  
                   405                                  410                                  415  
 Leu Phe Pro Val Val Trp Ser Ile Arg His Leu Gln Glu Ala Ser Ala  
                   420                                  425                                  430  
 Thr Asp Gly Lys Leu Pro Leu Thr  
                   435                                  440

## (2) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1838 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cdNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW051

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 114..1505

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1411..1603
- (D) OTHER INFORMATION: /label= SAC\_24170



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GTGACGGCT GGGTTGAAAA GTGACAACGG CCGGTGGATT TTAGGAGTTT GCTCGGTTTG	60
TAAGTCTCTT TTGGTGAGCT ACTGGGACTG CAGACTAGGA GGAGACTCCC AAA ATG	116
	Met
	1
GAA ACT CTG TCC TTC CCC AGA TAC AAC ATA GCT GAG ATT GTA GTT CAT	164
Glu Thr Leu Ser Phe Pro Arg Tyr Asn Ile Ala Glu Ile Val Val His	
	5 10 15
ATT CGC AAT AAA CTG TTA ACT GGA GCG GAT GGC AAA AAC CTC TCC AAG	212
Ile Arg Asn Lys Leu Leu Thr Gly Ala Asp Gly Lys Asn Leu Ser Lys	
	20 25 30
AGC GAT TTT CTT CCA AAC CCG AAG CCT GAA GTC CTG TAC ATG ATT TAC	260
Ser Asp Phe Leu Pro Asn Pro Lys Pro Glu Val Leu Tyr Met Ile Tyr	
	35 40 45
ATG AGA GCC TTA CAG TTA GTG TAT GGG GTC CGG CTG GAG CAT TTC TAC	308
Met Arg Ala Leu Gln Leu Val Tyr Gly Val Arg Leu Glu His Phe Tyr	
	50 55 60 65
ATG ATG CCG GTG AAC ATA GAA GTC ATG TAT CCA CAT ATA ATG GAG GGC	356
Met Met Pro Val Asn Ile Glu Val Met Tyr Pro His Ile Met Glu Gly	
	70 75 80
TTC TTA CCG GTC AGC AAT TTG TTC TTC CAC CTG GAC TCG TTT ATG CCC	404
Phe Leu Pro Val Ser Asn Leu Phe Phe His Leu Asp Ser Phe Met Pro	
	85 90 95
ATT TGC CGG GTG AAT GAC TTT GAG ATC GCC GAT ATT CTT TAT CCA AAA	452
Ile Cys Arg Val Asn Asp Phe Glu Ile Ala Asp Ile Leu Tyr Pro Lys	
	100 105 110
GCA AAC CGG ACA AGT CGT TTT TTA AGT GGC ATT ATC AAC TTT ATT CAC	500
Ala Asn Arg Thr Ser Arg Phe Leu Ser Gly Ile Ile Asn Phe Ile His	
	115 120 125
TTC AGA GAA ACA TGC CTG GAG AAG TAT GAA GAA TTT CTT TTG CAA AAT	548
Phe Arg Glu Thr Cys Leu Glu Lys Tyr Glu Glu Phe Leu Leu Gln Asn	
	130 135 140 145
AAA TCC TCT GTG GAC AAA ATA CAG CAG TTA AGC AAT GCA CAC CAG GAA	596
Lys Ser Ser Val Asp Lys Ile Gln Gln Leu Ser Asn Ala His Gln Glu	
	150 155 160
GCA TTG ATG AAA CTG GAA AAA CTC AAT TCG GTT CCC GTG GAG GAG CAG	644
Ala Leu Met Lys Leu Glu Lys Leu Asn Ser Val Pro Val Glu Glu Gln	
	165 170 175
GAA GAG TTC AAA CAG CTG AAG GAT GAC ATC CAG GAG CTG CAG CAC TTG	692
Glu Glu Phe Lys Gln Leu Lys Asp Asp Ile Gln Glu Leu Gln His Leu	
	180 185 190
CTG AAT CAA GAC TTC AGA CAG AAA ACG ACA CTG CTG CAG GAG AGA TAT	740
Leu Asn Gln Asp Phe Arg Gln Lys Thr Thr Leu Leu Gln Glu Arg Tyr	
	195 200 205
ACC AAA ATG AAA TCA GAT TTT TCA GAG AAA ACC AAG CAT GTT AAT GAG	788

Thr	Lys	Met	Lys	Ser	Asp	Phe	Ser	Glu	Lys	Thr	Lys	His	Val	Asn	Glu	
210					215					220					225	
CTA	AAG	TTG	TCA	GTA	GTT	TCT	TTG	AAA	GAA	GTT	CAA	GAC	AGT	TTG	AAA	836
Leu	Lys	Leu	Ser	Val	Val	Ser	Leu	Lys	Glu	Val	Gln	Asp	Ser	Leu	Lys	
				230					235					240		
AGC	AAA	ATT	GTG	GAT	TCT	CCA	GAG	AAG	CTG	AAG	AAC	TAT	AAA	GAG	AAG	884
Ser	Lys	Ile	Val	Asp	Ser	Pro	Glu	Lys	Leu	Lys	Asn	Tyr	Lys	Glu	Lys	
			245					250					255			
ATG	AAG	GAC	ACC	GTC	CAG	AAG	CTC	CGC	AGT	GCC	AGG	GAA	GAA	GTG	ATG	932
Met	Lys	Asp	Thr	Val	Gln	Lys	Leu	Arg	Ser	Ala	Arg	Glu	Glu	Val	Met	
		260					265					270				
GAG	AAG	TAT	GAT	ATC	TAT	AGA	GAT	TCT	GTG	GAT	TGC	TTG	CCT	TCC	TGT	980
Glu	Lys	Tyr	Asp	Ile	Tyr	Arg	Asp	Ser	Val	Asp	Cys	Leu	Pro	Ser	Cys	
	275					280					285					
CAG	CTG	GAG	GTG	CAG	TTA	TAT	CAA	AAG	AAA	TCA	CAG	GAC	CTT	GCA	GAT	1028
Gln	Leu	Glu	Val	Gln	Leu	Tyr	Gln	Lys	Lys	Ser	Gln	Asp	Leu	Ala	Asp	
290					295					300					305	
AAT	AGG	GAG	AAA	CTA	AGC	AGT	ATC	TTA	AAG	GAG	AGC	CTG	AAC	CTG	GAG	1076
Asn	Arg	Glu	Lys	Leu	Ser	Ser	Ile	Leu	Lys	Glu	Ser	Leu	Asn	Leu	Glu	
				310					315					320		
GGC	CAG	ATT	GAT	AGT	GAT	TCA	TCA	GAA	CTA	AAG	AAA	CTG	AAG	ACT	GAA	1124
Gly	Gln	Ile	Asp	Ser	Asp	Ser	Ser	Glu	Leu	Lys	Lys	Leu	Lys	Thr	Glu	
			325					330					335			
GAG	AAC	TCC	CTC	ATA	AGA	CTG	ATG	ACT	CTA	AAG	AAG	GAG	AGA	CTT	GCT	1172
Glu	Asn	Ser	Leu	Ile	Arg	Leu	Met	Thr	Leu	Lys	Lys	Glu	Arg	Leu	Ala	
		340					345					350				
ACC	ATG	CAG	TTC	AAA	ATA	AAC	AAG	AAG	CAG	GAG	GAT	GTG	AAA	CAG	TAC	1220
Thr	Met	Gln	Phe	Lys	Ile	Asn	Lys	Lys	Gln	Glu	Asp	Val	Lys	Gln	Tyr	
	355					360					365					
AAG	CGG	ACC	ATG	ATT	GAA	GAT	TGC	AAT	AAA	GTT	CAA	GAA	AAA	AGA	GAT	1268
Lys	Arg	Thr	Met	Ile	Glu	Asp	Cys	Asn	Lys	Val	Gln	Glu	Lys	Arg	Asp	
370					375					380					385	
GCT	GTC	TGC	GAG	CAA	GTA	ACC	GCC	ATT	AAT	CAA	GAC	ATC	CAC	AAG	ATT	1316
Ala	Val	Cys	Glu	Gln	Val	Thr	Ala	Ile	Asn	Gln	Asp	Ile	His	Lys	Ile	
				390				395						400		
AAA	TCT	GGG	ATT	CAG	CAG	CTA	AGA	GAC	GCC	GAA	AAA	CGG	GAG	AAA	CTG	1364
Lys	Ser	Gly	Ile	Gln	Gln	Leu	Arg	Asp	Ala	Glu	Lys	Arg	Glu	Lys	Leu	
			405					410					415			
AAG	TCT	CAG	GAA	ATC	TTG	GTA	GAC	TTG	AAA	AGT	GCT	TTG	GAG	AAG	TAC	1412
Lys	Ser	Gln	Glu	Ile	Leu	Val	Asp	Leu	Lys	Ser	Ala	Leu	Glu	Lys	Tyr	
		420					425					430				
CAT	GAG	GGC	ATC	GAG	AAG	ACG	ACG	GAG	GAG	TGC	TGC	ACT	AGA	ATA	GGA	1460
His	Glu	Gly	Ile	Glu	Lys	Thr	Thr	Glu	Glu	Cys	Cys	Thr	Arg	Ile	Gly	
		435				440						445				
GGG	AAG	ACT	GCC	GAG	CTG	AAG	AGG	AGG	ATG	TTC	AAA	ATG	CCG	CCC		1505
Gly	Lys	Thr	Ala	Glu	Leu	Lys	Arg	Arg	Met	Phe	Lys	Met	Pro	Pro		
450					455					460						

TGATCAACAG CCACCCGAAA ATGGCCTTTC GCTTTCTGTT TGGAGTAGTT ATATTGAAGC	1565
TAATAGAAGG ACCGCAGTCT CAGCTAACTA GCGTGGGTAC CATTGGTTCT CTGTCCTTTT	1625
ATGACCATGT GCTCCCTGTG TTTTCTTCTT GGTGATGGGG ATACAACTCA GGGCCTTGCA	1685
GGCTACACTG ACTTGCCTCC CTAGGCTCTA ATGTACCATG TACTATGTAG GCTTTTGCTA	1745
CAATTAAAGT AACGTGTACA GCTTTTATGT CCCTACTCTG TCTCCTTTTG TATGTGCTGG	1805
TTGGAATAAA CAAATAGTTA CTGACGTCAA AAA	1838

## (2) INFORMATION FOR SEQ ID NO:45:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 464 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met	Glu	Thr	Leu	Ser	Phe	Pro	Arg	Tyr	Asn	Ile	Ala	Glu	Ile	Val	Val	1	5	10	15
His	Ile	Arg	Asn	Lys	Leu	Leu	Thr	Gly	Ala	Asp	Gly	Lys	Asn	Leu	Ser	20	25	30	
Lys	Ser	Asp	Phe	Leu	Pro	Asn	Pro	Lys	Pro	Glu	Val	Leu	Tyr	Met	Ile	35	40	45	
Tyr	Met	Arg	Ala	Leu	Gln	Leu	Val	Tyr	Gly	Val	Arg	Leu	Glu	His	Phe	50	55	60	
Tyr	Met	Met	Pro	Val	Asn	Ile	Glu	Val	Met	Tyr	Pro	His	Ile	Met	Glu	65	70	75	80
Gly	Phe	Leu	Pro	Val	Ser	Asn	Leu	Phe	Phe	His	Leu	Asp	Ser	Phe	Met	85	90	95	
Pro	Ile	Cys	Arg	Val	Asn	Asp	Phe	Glu	Ile	Ala	Asp	Ile	Leu	Tyr	Pro	100	105	110	
Lys	Ala	Asn	Arg	Thr	Ser	Arg	Phe	Leu	Ser	Gly	Ile	Ile	Asn	Phe	Ile	115	120	125	
His	Phe	Arg	Glu	Thr	Cys	Leu	Glu	Lys	Tyr	Glu	Glu	Phe	Leu	Leu	Gln	130	135	140	
Asn	Lys	Ser	Ser	Val	Asp	Lys	Ile	Gln	Gln	Leu	Ser	Asn	Ala	His	Gln	145	150	155	160
Glu	Ala	Leu	Met	Lys	Leu	Glu	Lys	Leu	Asn	Ser	Val	Pro	Val	Glu	Glu	165	170	175	
Gln	Glu	Glu	Phe	Lys	Gln	Leu	Lys	Asp	Asp	Ile	Gln	Glu	Leu	Gln	His	180	185	190	
Leu	Leu	Asn	Gln	Asp	Phe	Arg	Gln	Lys	Thr	Thr	Leu	Leu	Gln	Glu	Arg	195	200	205	

Tyr Thr Lys Met Lys Ser Asp Phe Ser Glu Lys Thr Lys His Val Asn  
210 215 220

Glu Leu Lys Leu Ser Val Val Ser Leu Lys Glu Val Gln Asp Ser Leu  
225 230 235 240

Lys Ser Lys Ile Val Asp Ser Pro Glu Lys Leu Lys Asn Tyr Lys Glu  
245 250 255

Lys Met Lys Asp Thr Val Gln Lys Leu Arg Ser Ala Arg Glu Glu Val  
260 265 270

Met Glu Lys Tyr Asp Ile Tyr Arg Asp Ser Val Asp Cys Leu Pro Ser  
275 280 285

Cys Gln Leu Glu Val Gln Leu Tyr Gln Lys Lys Ser Gln Asp Leu Ala  
290 295 300

Asp Asn Arg Glu Lys Leu Ser Ser Ile Leu Lys Glu Ser Leu Asn Leu  
305 310 315 320

Glu Gly Gln Ile Asp Ser Asp Ser Ser Glu Leu Lys Lys Leu Lys Thr  
325 330 335

Glu Glu Asn Ser Leu Ile Arg Leu Met Thr Leu Lys Lys Glu Arg Leu  
340 345 350

Ala Thr Met Gln Phe Lys Ile Asn Lys Lys Gln Glu Asp Val Lys Gln  
355 360 365

Tyr Lys Arg Thr Met Ile Glu Asp Cys Asn Lys Val Gln Glu Lys Arg  
370 375 380

Asp Ala Val Cys Glu Gln Val Thr Ala Ile Asn Gln Asp Ile His Lys  
385 390 395 400

Ile Lys Ser Gly Ile Gln Gln Leu Arg Asp Ala Glu Lys Arg Glu Lys  
405 410 415

Leu Lys Ser Gln Glu Ile Leu Val Asp Leu Lys Ser Ala Leu Glu Lys  
420 425 430

Tyr His Glu Gly Ile Glu Lys Thr Thr Glu Glu Cys Cys Thr Arg Ile  
435 440 445

Gly Gly Lys Thr Ala Glu Leu Lys Arg Arg Met Phe Lys Met Pro Pro  
450 455 460

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1483 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW052

## (ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 247..765

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: complement (660..1062)

(D) OTHER INFORMATION: /label= SAC\_24882

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GTCGACCAAA AACAGGACAG ATTGCAACAG CCTCTACACT TCCTTCATGG CCTGCTCTTG	60
GGCCGTGGCG ACAAGCAGAA AGAAGCTTGA TTGTGGAGAG GTGCCCTAGC GGAGGGATGT	120
CTGCCACAGT CCAGGCACCC CACAGAGGGT AGGAGCTCCA TGTACGAGTG GCTTTGCATT	180
TCTTCAGAAA AGAATAGCTG ATGTGCAATC TCAGCCTTTC TGCAGAAGGA TTTATCAGGG	240
GCCAGC ATG GGC TCT GTG AAT TCT AGA GGC CAT AAG GCC CAG GTG GTA	288
Met Gly Ser Val Asn Ser Arg Gly His Lys Ala Gln Val Val	
1 5 10	
ATG CTG GGC CTC GAC TGT GCT GGC AAG ACC ACA ATC CTG TAC AAA CTG	336
Met Leu Gly Leu Asp Cys Ala Gly Lys Thr Thr Ile Leu Tyr Lys Leu	
15 20 25 30	
AAA GGA AAC CGG CTG GTG GAT ACC CTA CCC ACT GTT GGT TTT AAT GTA	384
Lys Gly Asn Arg Leu Val Asp Thr Leu Pro Thr Val Gly Phe Asn Val	
35 40 45	
GAG CCT CTT GAG GCT CCT GGA CAT GTG TCG CTG ACT CTC TGG GAC ATT	432
Glu Pro Leu Glu Ala Pro Gly His Val Ser Leu Thr Leu Trp Asp Ile	
50 55 60	
GGG GGA CAG ACC CAG CTC AGG GCT ACG TGG AAG GAC TAC CTG GAA GGC	480
Gly Gly Gln Thr Gln Leu Arg Ala Thr Trp Lys Asp Tyr Leu Glu Gly	
65 70 75	
ATT GAC CTC CTT GTG TAC GTG CTG GAC AGC ACA GAT GAA GCC CGC TTG	528
Ile Asp Leu Leu Val Tyr Val Leu Asp Ser Thr Asp Glu Ala Arg Leu	
80 85 90	
CCC GAG GCA GTG GCT GAG CTC GAG GAA GTC CTA GAA GAC CCC AAC ATG	576
Pro Glu Ala Val Ala Glu Leu Glu Glu Val Leu Glu Asp Pro Asn Met	
95 100 105 110	
GCT GGT GTT CCT TTC TTG GTA CTG GCC AAC AAG CAG GAG GCT CCT GAT	624
Ala Gly Val Pro Phe Leu Val Leu Ala Asn Lys Gln Glu Ala Pro Asp	
115 120 125	
GCT CTA CCA TTG CTT GAA ATC AGA AAC AGG CTG GAC CTG GAA AGG TTC	672
Ala Leu Pro Leu Leu Glu Ile Arg Asn Arg Leu Asp Leu Glu Arg Phe	
130 135 140	
CAA GAC CAT TGC TGG GAG CTG CGG GCT TGC AGC GCT CTC ACA GGT CAG	720
Gln Asp His Cys Trp Glu Leu Arg Ala Cys Ser Ala Leu Thr Gly Gln	
145 150 155	
GGG CTA CAG GAA GCC CGG CAG AGC CTG CTG CAT TTG CTG AGA TCC	765

Gly	Leu	Gln	Glu	Ala	Arg	Gln	Ser	Leu	Leu	His	Leu	Leu	Arg	Ser	
160						165					170				
TGATGAAATG	CAGTAACCTC	TAACCTCCAGG	CTGGAGCAGG	ACAGGGAGCA	GGGTCAGCCA										825
GACCCCTGA	GCAGGCAAAT	AGATTATCCT	TGCTTCTAAG	AACAGGAGGG	GCCAGCTGAT										885
CCTTGAGAAA	GTGAAGCTCC	GTTTGCCACA	TGAGAAACCA	CGTTTCTCTT	GGACGGTGTG										945
TTTTCTTCTC	AGGACTTCTT	CTCAGTGGCG	GTTATGCGGA	AAATCAACTG	TTCTTCAAGT										1005
AACAATCGTC	AGAGGGAATA	TAAAAACCT	CAGAATACCC	CCATGTAACA	GCTCAGTACC										1065
AAGTGTCAAA	GTCACAAACT	TGGTGACATA	ATTATTGATG	ACCAAGTTCA	CAAATGAAAA										1125
GAGGTTTCAA	AACTACATCT	GAGTGTCACT	GGGCCTATCG	CTGTTATCCG	AGTGTCACTG										1185
GGCCTATCAT	TGTCATCCGA	GTGTCACCAG	GCCTATCACG	GTCATCTGCC	TGTCACTGGG										1245
CCTATCACAG	TCATCAGGAG	CGACTCTTCC	TACTGGCCTT	GGGGCAGAAG	CCACTGTCAC										1305
TGGTGGGCTC	ACAGCAGGAG	GGGCAGGTCT	GCTCTCATAA	TGCTGCAGAA	GGCAGAAGTC										1365
CATGGGGCCC	TAGATCCTCC	CAGAATTCCC	CCTGATCTCC	AACTCTGACT	AACTCTTTTC										1425
TGCCCTATAT	CCTCTGTTGT	AAAAAAGTAA	AATAAACTC	CTCTTATACT	CTCAAAAA										1483

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 173 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met	Gly	Ser	Val	Asn	Ser	Arg	Gly	His	Lys	Ala	Gln	Val	Val	Met	Leu
1				5					10					15	
Gly	Leu	Asp	Cys	Ala	Gly	Lys	Thr	Thr	Ile	Leu	Tyr	Lys	Leu	Lys	Gly
		20					25						30		
Asn	Arg	Leu	Val	Asp	Thr	Leu	Pro	Thr	Val	Gly	Phe	Asn	Val	Glu	Pro
		35				40					45				
Leu	Glu	Ala	Pro	Gly	His	Val	Ser	Leu	Thr	Leu	Trp	Asp	Ile	Gly	Gly
	50				55					60					
Gln	Thr	Gln	Leu	Arg	Ala	Thr	Trp	Lys	Asp	Tyr	Leu	Glu	Gly	Ile	Asp
65				70					75					80	
Leu	Leu	Val	Tyr	Val	Leu	Asp	Ser	Thr	Asp	Glu	Ala	Arg	Leu	Pro	Glu
			85					90						95	
Ala	Val	Ala	Glu	Leu	Glu	Glu	Val	Leu	Glu	Asp	Pro	Asn	Met	Ala	Gly
		100					105						110		
Val	Pro	Phe	Leu	Val	Leu	Ala	Asn	Lys	Gln	Glu	Ala	Pro	Asp	Ala	Leu
		115					120						125		

Pro Leu Leu Glu Ile Arg Asn Arg Leu Asp Leu Glu Arg Phe Gln Asp  
 130 135 140  
 His Cys Trp Glu Leu Arg Ala Cys Ser Ala Leu Thr Gly Gln Gly Leu  
 145 150 155 160  
 Gln Glu Ala Arg Gln Ser Leu Leu His Leu Leu Arg Ser  
 165 170

## (2) INFORMATION FOR SEQ ID NO:48:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2288 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW054
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 1952..2152
  - (D) OTHER INFORMATION: /label= SAC\_24183

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GTCGACGGTG GGGTCTAACT GAGGTCTTGT ATTTGGTTTT TCCTGAAGAT CGACCTCTGG	60
CATCAGGAGG CTTGAGCAAG GAGAAGAGTG GATGATGAAG GAAAGAGACA AGGAAAGAAA	120
GAGAAGTGTG CACAAAGAAA CTTGTATTAT TATTAATTAG CACCTAGCTT GTTTGTGTCT	180
GATACACCAC CAAGTAGTAA TTGTTGAAAA AAGCGAAGAA GAAAAAAAAA AAAACAAAAA	240
AACCAAACAG TGGGTACTCA AATAAGATAG GAGAAAAATG AAGAGAACAG ACCCAGTTCT	300
CGACCCTTGC TTCTCGAAGG TCCTCCCACC AGGCTGCCAA AGCAAGATGG TGTGCTCTG	360
ATCCAGTCAG TATTCCTTTG ACTTTTTTTT TTAATCTCCA GGTTTGGTT CAGGCTCCCA	420
TATTCATACC CTGGCTCATT TAGCTTTCCC TCATGTTGTG GGTCTTCTG TCCCTCACCC	480
CCTTACTCTC CCCACTGATA TTCTTCCCAC GTCAAGACTG TGGCTCTGGA AGAAATATCC	540
ACCATTGCA GAGCTGATGT TCTGTAGATC GTAATGTTGA AGCGCTGGGT GTCCTGGTTG	600
GCAGAATCAC TCCTGTATTA CTCTGGTACA TAGGTGTCCT CTGNTAGACT CCCTGGCCTT	660
AGTCATGGGG TGTTTTCTAG AGGCAGACTA AGACAGGAGT CAAAAAGAT TTAGAGGAAG	720
GAGCTGAGGA AAGAAAGACA GTTGTGGGAG GAAATCAAG TTCTACTCAG GATCCCGAGT	780
GTTTCTGTAG ATGTAGATTG GAATGTGTCC ATAACAGAGA GGCCAGTGAG AGACATCCCC	840

AAGGACCTGC CAGGCTTTCC TTCGCTCCAG GAAGACGCAC CATCACTCAA AAGGGGTTC	900
CTAGAAAGAA AGACAAGTGA CTTAAAAAAT CTGCCAGTGG GTTCTTGAAG TCATCGAACC	960
TAGGAGGTAG AAAGCTCAGG GGGGTCATGT TATCGAAGGG AGGTGAGGTT CAGTCAGGCT	1020
AACTATTAGG GAGANTGTCA AAGGAGGCTA CAGAAATGGG ATTCTGGCTG TGGGGCATT	1080
AAGGCTATCA AGTCAACAGG TAAGAGTGA GTGGCTCAGC AGCTTCTCTG GCGCTGAAAG	1140
TGGTGTCTCT TAACAGTTAA TTCCCTTCCC TTTTCCATCA CAGAAGGAAT CATTCCAGCC	1200
CTCCCTGGGA TTATGGTCCC CCAGTGACCA TGTTCATCT GCTGGATTTT TGATCAGTTT	1260
AGAAGGAAAA AAACACTTGG GTTCGGTGGT ATTTGCTTAT GACCTTAGCA CTTTGAGCT	1320
GAGAAGAGGA TTGCCATAGT TTCCAGACCA GTCTGGTTTA CATAGTGAGT TCTAGGGCTG	1380
CGTAATGAGA TGTCTCAAGA ATGGAGAGGG GCGTTAGGGA GAACTTACTT AGCTCTTCAA	1440
GAAACTGTCC TGTAAACAAA GAGAAATGAG TCGATGGGTG ACCAGTGGGA GCCTNCCCCC	1500
CACCTCCTCA GACAACCCTT CTTTCCATT AAGGGGCTG TGCATCTCAG TAAGTGGTCC	1560
TCAGTTTCC TGAGGATTCT GATGGTGAGA GTAGAGAGAG CAAAAGATT CTCAGCCCTG	1620
GGATCTAGAA CCTTGGTTCC ACTCTTTTCC CGTTCTGGTC TGCCTGAGCT CCCCCCCCC	1680
GCCCCCAGT GGTCTGTTT GCTGTCCAGA CATTGCACCT GCCAGTCTT GGTGCAGCCA	1740
GTGTAGCAG ATGATCAAAG CAACCAGNGG TCCAAGAACA GGGACAATCA ACAAACAGTG	1800
TTATCTTCCA GCAGGGCACT CTCAGGAAGT GAGGATCTGA ATGGTAAAGC AGAACAGAGT	1860
GTTCCTCT ATGCCACAGC CGCTCATGCC ACACCCACTT GAAGCCTCCA AAGCCAGNAC	1920
AAGCTGTAGA GTCCATGAAT GTAGATGAGG TACATTGTGC GCCTTGCTTT TACTTCCTTT	1980
AGCTTCTGAT TGAATTTAGC TGTATCTTTT TAACCGATTT ATCGTAGAAN ACAATGGGTG	2040
GTGTGACCCT NTCATGCTTT GTTGATCTTC AGCCACACGC TGCCCTCCCT CGCCCCACT	2100
TCTAGCATTT GTGTGTGTGC GCGTGTGTTT TTGAAGAGGG GTTNTTGCTT GTACAATCCC	2160
TGGCTGGCCT TGAACCTGAG AAGTCCTCCT GCCTCCGCCT TCCAGGTGCT GGGATTACAA	2220
GTGTCAACC CTGCTTTCAC TAATGTTTCC TTTTATTAA AGATTATTT ATTATTTAAA	2280
AAGTCGAC	2288

## (2) INFORMATION FOR SEQ ID NO:49:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2349 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO



(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW055

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 214..1329

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1953..2065  
(D) OTHER INFORMATION: /label= SAC\_24449

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GTCGACAATG CTTTCAAGCA AACAGCCGCT ACCAGAACAA CTCTGTGACT AGGAGAAAGA	60
GGCCAAGGCG CGAGTTCCCA AACCCCTGAAG CCACAGGAGC CCAAGGGGGA GACTGGAGAG	120
AAGGTATTTC TAGCCGTGCA CAGTGACTGA TCAGCCGAGG ACCTACCCTC GAGTTCTATT	180
CTGCAACAGT GAACAGTGAA GGTCTCCAAA GAG ATG GAG TAC GAG GGT TAC AAC	234
Met Glu Tyr Glu Gly Tyr Asn	
1 5	
GAC TCC AGC ATC TAC GGT GAG GAG TAT TCT GAC GGC TCG GAC TAC ATC	282
Asp Ser Ser Ile Tyr Gly Glu Glu Tyr Ser Asp Gly Ser Asp Tyr Ile	
10 15 20	
GTG GAC TTG GAG GAG GCG GGT CCA CTG GAG GCC AAG GTG GCC GAG GTC	330
Val Asp Leu Glu Glu Ala Gly Pro Leu Glu Ala Lys Val Ala Glu Val	
25 30 35	
TTC CTG GTG GTA ATC TAC AGC TTG GTG TGC TTC CTC GGG ATC CTA GGC	378
Phe Leu Val Val Ile Tyr Ser Leu Val Cys Phe Leu Gly Ile Leu Gly	
40 45 50 55	
AAT GGC CTG GTG ATT GTC ATC GCC ACC TTC AAG ATG AAG AAG ACG GTG	426
Asn Gly Leu Val Ile Val Ile Ala Thr Phe Lys Met Lys Lys Thr Val	
60 65 70	
AAC ACC GTG TGG TTT GTC AAC CTG GCC GTG GCT GAC TTC CTG TTC AAC	474
Asn Thr Val Trp Phe Val Asn Leu Ala Val Ala Asp Phe Leu Phe Asn	
75 80 85	
ATC TTC TTG CCC ATC CAC ATC ACC TAT GCC GCT ATG GAC TAC CAC TGG	522
Ile Phe Leu Pro Ile His Ile Thr Tyr Ala Ala Met Asp Tyr His Trp	
90 95 100	
GTG TTC GGG AAA GCC ATG TGC AAG ATT AGT AGC TTT CTG CTA AGC CAC	570
Val Phe Gly Lys Ala Met Cys Lys Ile Ser Ser Phe Leu Leu Ser His	
105 110 115	
AAC ATG TAC ACC AGC GTC TTC CTG CTC ACT GTC ATC AGC TTC GAC CGC	618
Asn Met Tyr Thr Ser Val Phe Leu Leu Thr Val Ile Ser Phe Asp Arg	
120 125 130 135	
TGC ATC TCC GTG CTC CTC CCC GTC TGG TCC CAG AAC CAC CGC AGC GTG	666
Cys Ile Ser Val Leu Leu Pro Val Trp Ser Gln Asn His Arg Ser Val	
140 145 150	
CGT CTG GCC TAC ATG ACC TGC GTG GTT GTC TGG GTC CTG GCT TTC TTC	714
Arg Leu Ala Tyr Met Thr Cys Val Val Val Trp Val Leu Ala Phe Phe	

155										160										165										
CTG	AGT	TCC	CCG	TCC	CTC	GTC	TTC	CGG	GAC	ACG	GTC	AGC	ACC	AGC	CAC															
Leu	Ser	Ser	Pro	Ser	Leu	Val	Phe	Arg	Asp	Thr	Val	Ser	Thr	Ser	His															
		170						175					180																	
GGG	AAG	ATA	ACC	TGC	TTC	AAC	AAC	TTC	AGC	CTG	GCC	GCG	CCC	GAG	CCT															
Gly	Lys	Ile	Thr	Cys	Phe	Asn	Asn	Phe	Ser	Leu	Ala	Ala	Pro	Glu	Pro															
		185					190					195																		
TTC	TCT	CAT	TCC	ACC	CAC	CCG	CGA	ACA	GAC	CCG	GTA	GGG	TAC	AGC	AGA															
Phe	Ser	His	Ser	Thr	His	Pro	Arg	Thr	Asp	Pro	Val	Gly	Tyr	Ser	Arg															
		200				205				210					215															
CAT	GTG	GCG	GTC	ACC	GTC	ACC	CGC	TTC	CTC	TGT	GGC	TTC	CTG	ATC	CCC															
His	Val	Ala	Val	Thr	Val	Thr	Arg	Phe	Leu	Cys	Gly	Phe	Leu	Ile	Pro															
				220					225					230																
GTC	TTC	ATC	ATC	ACG	GCC	TGT	TAC	CTC	ACC	ATC	GTC	TTC	AAG	TTG	CAG															
Val	Phe	Ile	Ile	Thr	Ala	Cys	Tyr	Leu	Thr	Ile	Val	Phe	Lys	Leu	Gln															
			235					240					245																	
CGC	AAC	CGC	CTG	GCC	AAG	ACC	AAG	AAG	CCC	TTC	AAG	ATC	ATC	ATC	ACC															
Arg	Asn	Arg	Leu	Ala	Lys	Thr	Lys	Lys	Pro	Phe	Lys	Ile	Ile	Ile	Thr															
		250					255					260																		
ATC	ATC	ATC	ACC	TTC	TTC	CTC	TGC	TGG	TGC	CCC	TAC	CAC	ACA	CTC	TAC															
Ile	Ile	Ile	Thr	Phe	Phe	Leu	Cys	Trp	Cys	Pro	Tyr	His	Thr	Leu	Tyr															
		265				270					275																			
CTG	CTG	GAG	CTC	CAC	CAC	ACG	GCT	GTG	CCA	GCC	TCT	GTC	TTC	AGC	CTG															
Leu	Leu	Glu	Leu	His	His	Thr	Ala	Val	Pro	Ala	Ser	Val	Phe	Ser	Leu															
		280			285				290						295															
GGA	CTG	CCC	CTG	GCC	ACA	GCC	GTT	GCC	ATC	GCC	AAC	AGC	TGT	ATG	AAC															
Gly	Leu	Pro	Leu	Ala	Thr	Ala	Val	Ala	Ile	Ala	Asn	Ser	Cys	Met	Asn															
				300				305						310																
CCC	ATC	CTG	TAC	GTC	TTC	ATG	GGC	CAC	GAC	TTC	AAA	AAA	TTC	AAG	GTG															
Pro	Ile	Leu	Tyr	Val	Phe	Met	Gly	His	Asp	Phe	Lys	Lys	Phe	Lys	Val															
			315				320						325																	
GCC	CTT	TTC	TCC	CGC	CTG	GTG	AAT	GCC	CTG	AGC	GAG	GAC	ACA	GGA	CCC															
Ala	Leu	Phe	Ser	Arg	Leu	Val	Asn	Ala	Leu	Ser	Glu	Asp	Thr	Gly	Pro															
		330				335						340																		
TCC	TCC	TAC	CCC	AGT	CAC	AGG	AGC	TTC	ACC	AAG	ATG	TCC	TCA	TTG	ATT															
Ser	Ser	Tyr	Pro	Ser	His	Arg	Ser	Phe	Thr	Lys	Met	Ser	Ser	Leu	Ile															
		345				350					355																			
GAG	AAG	GCT	TCA	GTG	AAT	GAG	AAA	GAG	ACC	AGC	ACC	CTC	TGAGCCTTCC																	
Glu	Lys	Ala	Ser	Val	Asn	Glu	Lys	Glu	Thr	Ser	Thr	Leu																		
		360			365			370																						
CTGGGACTGT	CCCCAATGGG	CTCCACAGCC	CAGGGACTTG	TCTCCTGAGG	TAGGAGACAC																									
GCCTGGAGCC	TTTGGGTATG	CCCCAGCGCC	CGCTACGTCT	TGCGCAAGGC	GGCCCATTTGT																									
TTTGGGTGGA	GTTCCCAAGT	GTGGACACTC	TTCCAGCAAA	ATGGCAGACA	AGCAACCTGA																									
GCCCCTGCAG	CAGGAGTATA	GGACCAACTT	CCACCGGCTC	AGAAAAGGGA	GGGTCTCTGA																									
AACCAAGGTC	TGAGCTGTGA	CCCACATACA	GGCACACAAT	TTCACGTGG	ATGCCCAAA																									

CATGCTGCAC AGTCTCCAGT GTGGGTGAGG ACGTCACTGC AGACCTGTTA TGGCGACATG 1699  
 ACAGTCAAAC CAGCAAATAC CCCACCACCA CTGTCATCCT CCAAGATCTT GACTTTGGAT 1759  
 TTCAGAAAAA TGGGGAGGAC CAGGACCTGA GGGGCTTCAT AGAACTTGCT TGCAAGGGCA 1819  
 GCCCAGGTGT GTGTGTGTGT ATGTGCTTGT GTGTGTGTGT GTGTGTATGT GCATGTGTGT 1879  
 GTGTGTGTGT GTGTGTGTGT GTGTGTGTTG AAGAGGTGGA GACAGAGAGC AGAGAACCCC 1939  
 AAAGGATGTG GGTACAGGAT ATCTCACACC ATCATGCCAC AGGGAGACAG AGCTACAACG 1999  
 ATGGGTGCTT GGGGGTTCAG ATGGTGGCAG CTCTTTGCTT CAGTGTGTAT ACCCCCCTAG 2059  
 CACTGTACCC CCAGGCTGGA GGGGGGCTCT TCACAGAAGC TTCCAACCCA TTCCAGCACT 2119  
 GGCTGGAAGG GCTTGGGATG TCCCTGGAGA CTGGTCTGGG ACTATCTTCA AAGCGTCTTC 2179  
 CAAAGGCTGC TGCTGACCAG CTCACCTGAC TCCTCCTTGA ACACACTCCC CACGTGGGGC 2239  
 ATCACAGCCC CGATGGCAGC CCAGGATCAG GCCTCCCTCC CTTGACAGGG AAATGACTCA 2299  
 CCCAACCCCA TGCACGTGCA GTAACAAAGA CCCAGGCCCA GCTCCAAAAA 2349

## (2) INFORMATION FOR SEQ ID NO:50:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Met Glu Tyr Glu Gly Tyr Asn Asp Ser Ser Ile Tyr Gly Glu Glu Tyr  
 1 5 10 15  
 Ser Asp Gly Ser Asp Tyr Ile Val Asp Leu Glu Glu Ala Gly Pro Leu  
 20 25 30  
 Glu Ala Lys Val Ala Glu Val Phe Leu Val Val Ile Tyr Ser Leu Val  
 35 40 45  
 Cys Phe Leu Gly Ile Leu Gly Asn Gly Leu Val Ile Val Ile Ala Thr  
 50 55 60  
 Phe Lys Met Lys Lys Thr Val Asn Thr Val Trp Phe Val Asn Leu Ala  
 65 70 75 80  
 Val Ala Asp Phe Leu Phe Asn Ile Phe Leu Pro Ile His Ile Thr Tyr  
 85 90 95  
 Ala Ala Met Asp Tyr His Trp Val Phe Gly Lys Ala Met Cys Lys Ile  
 100 105 110  
 Ser Ser Phe Leu Leu Ser His Asn Met Tyr Thr Ser Val Phe Leu Leu  
 115 120 125  
 Thr Val Ile Ser Phe Asp Arg Cys Ile Ser Val Leu Leu Pro Val Trp  
 130 135 140

Ser Gln Asn His Arg Ser Val Arg Leu Ala Tyr Met Thr Cys Val Val  
 145 150 155 160  
 Val Trp Val Leu Ala Phe Phe Leu Ser Ser Pro Ser Leu Val Phe Arg  
 165 170 175  
 Asp Thr Val Ser Thr Ser His Gly Lys Ile Thr Cys Phe Asn Asn Phe  
 180 185 190  
 Ser Leu Ala Ala Pro Glu Pro Phe Ser His Ser Thr His Pro Arg Thr  
 195 200 205  
 Asp Pro Val Gly Tyr Ser Arg His Val Ala Val Thr Val Thr Arg Phe  
 210 215 220  
 Leu Cys Gly Phe Leu Ile Pro Val Phe Ile Ile Thr Ala Cys Tyr Leu  
 225 230 235 240  
 Thr Ile Val Phe Lys Leu Gln Arg Asn Arg Leu Ala Lys Thr Lys Lys  
 245 250 255  
 Pro Phe Lys Ile Ile Ile Thr Ile Ile Ile Thr Phe Phe Leu Cys Trp  
 260 265 270  
 Cys Pro Tyr His Thr Leu Tyr Leu Leu Glu Leu His His Thr Ala Val  
 275 280 285  
 Pro Ala Ser Val Phe Ser Leu Gly Leu Pro Leu Ala Thr Ala Val Ala  
 290 295 300  
 Ile Ala Asn Ser Cys Met Asn Pro Ile Leu Tyr Val Phe Met Gly His  
 305 310 315 320  
 Asp Phe Lys Lys Phe Lys Val Ala Leu Phe Ser Arg Leu Val Asn Ala  
 325 330 335  
 Leu Ser Glu Asp Thr Gly Pro Ser Ser Tyr Pro Ser His Arg Ser Phe  
 340 345 350  
 Thr Lys Met Ser Ser Leu Ile Glu Lys Ala Ser Val Asn Glu Lys Glu  
 355 360 365  
 Thr Ser Thr Leu  
 370

## (2) INFORMATION FOR SEQ ID NO:51:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2577 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW056
- (ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 8..523

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (738..953)  
(D) OTHER INFORMATION: /label= SAC\_24326

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTCGACG	GGC	TAC	CTT	AGC	CAG	CTG	ATG	CTT	CCA	GTG	ACA	GAG	CCA	CTT	49
	Gly	Tyr	Leu	Ser	Gln	Leu	Met	Leu	Pro	Val	Thr	Glu	Pro	Leu	
	1				5					10					
TGT	CCT	TTG	CAC	AGC	CTG	ACA	CCC	TAT	CAG	GTC	CCC	TTC	AAT	GCA	GTT
Cys	Pro	Leu	His	Ser	Leu	Thr	Pro	Tyr	Gln	Val	Pro	Phe	Asn	Ala	Val
15					20					25					30
GCT	ATC	CGG	GTC	ACG	CAC	GCC	GAC	GTT	GCA	CCT	ACC	CAC	ATA	CTG	TAC
Ala	Ile	Arg	Val	Thr	His	Ala	Asp	Val	Ala	Pro	Thr	His	Ile	Leu	Tyr
				35					40					45	
GCT	ATG	AAT	GCC	AGC	TGG	GTC	GGC	CTT	TGC	AAG	ATT	GTG	GAT	GAT	ATG
Ala	Met	Asn	Ala	Ser	Trp	Val	Gly	Leu	Cys	Lys	Ile	Val	Asp	Asp	Met
			50					55					60		
AAA	GGC	TAC	ACT	CGG	GGC	CCC	ATT	CTG	CTG	GCC	CAG	AAC	CCC	ATA	TGT
Lys	Gly	Tyr	Thr	Arg	Gly	Pro	Ile	Leu	Leu	Ala	Gln	Asn	Pro	Ile	Cys
		65					70					75			
GAC	TGT	TTG	GGC	TTT	GGT	ATC	TGC	AGA	GGC	ATT	GAC	ATG	GAC	AAG	CGG
Asp	Cys	Leu	Gly	Phe	Gly	Ile	Cys	Arg	Gly	Ile	Asp	Met	Asp	Lys	Arg
	80					85					90				
CTG	TAC	CAC	ATC	CTC	ACC	CCT	CTA	CCA	CCA	GAG	GAG	TTA	AGA	ACT	GTG
Leu	Tyr	His	Ile	Leu	Thr	Pro	Leu	Pro	Pro	Glu	Glu	Leu	Arg	Thr	Val
95					100					105					110
AAC	TGT	CTG	CTG	GTT	GGC	ACC	ATT	TCC	ATT	CCA	CAT	TGT	ATT	TTC	AAG
Asn	Cys	Leu	Leu	Val	Gly	Thr	Ile	Ser	Ile	Pro	His	Cys	Ile	Phe	Lys
				115					120					125	
AAT	CAG	CCT	GGG	ACT	GAA	GGG	ACA	GTT	CCT	TAT	GTC	ACC	AGG	GAT	TAT
Asn	Gln	Pro	Gly	Thr	Glu	Gly	Thr	Val	Pro	Tyr	Val	Thr	Arg	Asp	Tyr
		130						135						140	
AAT	TTA	AAA	CTT	CTT	GGA	GCA	TCA	GAG	AAA	ATT	GGG	GAG	AGA	GAG	TAT
Asn	Leu	Lys	Leu	Leu	Gly	Ala	Ser	Glu	Lys	Ile	Gly	Glu	Arg	Glu	Tyr
		145					150					155			
AGA	AAT	ATA	CTA	CCT	AGA	CAC	AAA	TCA	CGG	CAA	AGA	AGA	AAA		
Arg	Asn	Ile	Leu	Pro	Arg	His	Lys	Ser	Arg	Gln	Arg	Arg	Lys		
	160					165					170				
TGAAGCCTGG	GTGTACAGAG	AGACTTTCGC	CTCGACACAT	TGTCCCCAAG	CCTAGAGTCA										583
GCAGACTGAG	CGGCTGTGT	GGCTGCTGAC	CTCGTGCCCA	TGCCCAGTGT	TCATATAACT										643
GGAAAAATGTG	CTGTGGTAAG	AGTCATTCAA	GCCCCCTTCGT	TTTTTTCCCC	TCTAAGACAT										703
GGAGGGCCTC	ATATAGCTCA	GACTGGCCTC	AGTACTCTGG	TCCTCCTGTC	TCACTGTCCA										763

AGGGCTGCGC ATCCGGCTTT TGGTCTCATA CTTGGCCCAG ACCTTGCCGA CTGGTTGTCC	823
TGGGTGTGTG GTGCTCCTGG AGAGCTGGGA AGCATTGAAG AAAGCTGTGG TGTCCCTCAA	883
GAAGACCAGG GCCACAAAAG TGTGCCTAGA AACTGAGTAG CAGGCTGACC ACAGGTTCTC	943
GGGTGCAGTG TACGGACTGG CAGGGACCTG AGCAGGCTCA GTGGCCGTGA GTAAGAGGCT	1003
TACTCCGTTG AACCTTGTGG ACCTTAGCAC CCCAGTGTGG CGGCTCATTG GCCTGATCCT	1063
TTGATGCTTG CCATCGAATC TCCCAATGCT GCAATCAAGT GTTAGAGGAC GGTACGTCGT	1123
CTTGTTAATT TCTGGGTCTT TAATCATGAG GCTTTATGTT GAAGAACTTG CTCTGACATT	1183
CAGAAAGCCT CACGTTAGTT AAAGAAATCC ATGGTTCCCT AATTCTGGGG ATGTTCTGAC	1243
AGACCTTGAG GCTGTAGATA CGTCAGAAAG CCTCCATGTT AGTTAAGGAA ATCCATGGTT	1303
CCCTAATTCT GGGGATGTTT TGACTGGGAA TGTCTTGACA GACCTTGAGG CTGAGATACT	1363
ACCATGGGCC CACTAGCCTG GGACAGGCAC ACATATGCCT GGCTTGGGCA GACCTTTGCT	1423
ATCAACGATT TCTTCTAAGG CCCCAGAGGC CACTGCAAGT GGATGTTGCA ATCCGGAGTG	1483
ACTGGGTCGG GTCATTTTGG TGGTTGCAAA GCAGGACCCT GAGGAGCCTC TTGGGCTGGC	1543
ATGATGGATC TGACAAGACT CCACCTTGCA TGAATCAGG AAGGACCAAG GCTATGGAAG	1603
GCTCACAACA TTGATTGTCT CCGTCTTGTC TGTAAGTACA GGAGTGTGGC CAGTCTCCTC	1663
TGCTTGGCGC CATCTTTGTG GTGGCCTCAG CATCTGAAAA TCCTCCGAGG AGGCAGAGCT	1723
GCCTCTTCTG GATTAGTCGG ATTCCTGTGA GATCACTCGG GCTCACAAC ACTCCCTGCC	1783
CCACCCCCAA CACCATCTCT GTCTCACTTG AGCTTGTGGT GAGCCTTGC AGAGGGGAGG	1843
AGCTCTTCTT CCCCAGGCAC TCTCTGGGCT CCGTATCCTC AAGCCTCTGC TTCCTGGGTG	1903
CTCAGATCAC AGGTGTGTGC CGTCTTAGCC TTGCGGACAG TTTTAAGACA GCCTTGGTCA	1963
CCTTCATCTC CCTCCCCGGG TCTTCCTGTC TTCACAATAA AGATCCCGAG TCCAAAAAAA	2023
AAAAAAAAAA AAATTCCAGG AGGAAATCTG GGTGCACAGA ACTTATCTGA GTCGGCTTTG	2083
GTGAAGGAGA TCCTCAAGGA GCAGGAGAAC AGGAAGGGCC TCATAGCTGC CATCTGTGCG	2143
GGTCCTACGG CCCTGCTGGC TCACGAAGTA GGCTTTGGAT GCAAGGTTAC ATCGCACCCA	2203
TTGGCTAAGG ACAAATGAT GAACGGCAGT CACTACAGCT ACTCAGAGAG CCGTGTGGAG	2263
AAGGACGGCC TCATCCTCAC CAGCCGTGGG CTTGGGACCA GCTTCGAGTT TGCGCTGGCC	2323
ATTGTGGAGG CACTCAGTGG CAAGGACATG GCTAACCAAG TGAAGGCCCC GCTTGTCTC	2383
AAAGACTAGA GCCCAAGCCC TGGACCCTGG ACCCCAGGC TGAGCAGGCA TTGGAAGCCC	2443
ACTAGTGTGT CCACAGCCCA GTGAACCTCA GGAACCTAACG TGTGAAGTAG CCCGCTGCTC	2503
AGGAATCTCG CCCTGGCTCT GTACTATCTT GAGCCTTGCT AGTAGAATAA ACAGTCCCC	2563
AAGCTCCTGA AAAA	2577

## (2) INFORMATION FOR SEQ ID NO:52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```

Gly Tyr Leu Ser Gln Leu Met Leu Pro Val Thr Glu Pro Leu Cys Pro
 1           5           10           15
Leu His Ser Leu Thr Pro Tyr Gln Val Pro Phe Asn Ala Val Ala Ile
          20           25           30
Arg Val Thr His Ala Asp Val Ala Pro Thr His Ile Leu Tyr Ala Met
          35           40           45
Asn Ala Ser Trp Val Gly Leu Cys Lys Ile Val Asp Asp Met Lys Gly
          50           55           60
Tyr Thr Arg Gly Pro Ile Leu Leu Ala Gln Asn Pro Ile Cys Asp Cys
          65           70           75           80
Leu Gly Phe Gly Ile Cys Arg Gly Ile Asp Met Asp Lys Arg Leu Tyr
          85           90           95
His Ile Leu Thr Pro Leu Pro Pro Glu Glu Leu Arg Thr Val Asn Cys
          100          105          110
Leu Leu Val Gly Thr Ile Ser Ile Pro His Cys Ile Phe Lys Asn Gln
          115          120          125
Pro Gly Thr Glu Gly Thr Val Pro Tyr Val Thr Arg Asp Tyr Asn Leu
          130          135          140
Lys Leu Leu Gly Ala Ser Glu Lys Ile Gly Glu Arg Glu Tyr Arg Asn
          145          150          155          160
Ile Leu Pro Arg His Lys Ser Arg Gln Arg Arg Lys
          165          170

```

## (2) INFORMATION FOR SEQ ID NO:53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1339 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW057

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..1023

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1032..1176

(D) OTHER INFORMATION: /label= SAC\_23926

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GTCGAC GGC GCG CTG GAG GAC GCA CGG CTT GGC GGG AAG GCG AGG GGA	48
Gly Ala Leu Glu Asp Ala Arg Leu Gly Gly Lys Ala Arg Gly	
1 5 10	
CCG TGC GGA GCC TCC GGT GCT GGC GGC GGC CAC CAG TTC GGG AGC CGG	96
Pro Cys Gly Ala Ser Gly Ala Gly Gly Gly His Gln Phe Gly Ser Arg	
15 20 25 30	
GAG CCG GCC GCA GCA CGC GGA ACA GAG CAG AGG GTG GCG GGG CCC GGC	144
Glu Pro Ala Ala Ala Arg Gly Thr Glu Gln Arg Val Ala Gly Pro Gly	
35 40 45	
GAA ACC ACG ATG AAG CCG AAC AGC GAG GAA GAG GAG GAG TTG GTG CAG	192
Glu Thr Thr Met Lys Pro Asn Ser Glu Glu Glu Glu Glu Leu Val Gln	
50 55 60	
GGC GTG GGC CCC TGG GAC GAG TGC TTC GAG GTG GCC GTG CAG TTG GCG	240
Gly Val Gly Pro Trp Asp Glu Cys Phe Glu Val Ala Val Gln Leu Ala	
65 70 75	
TTG CGT GCG GGA CAA ATC ATC AGA AAG GCC CTC ACT GAG GAA AAA CAC	288
Leu Arg Ala Gly Gln Ile Ile Arg Lys Ala Leu Thr Glu Glu Lys His	
80 85 90	
GTC TCG ACG AAA ACA TCT GCT GCA GAT CTT GTG ACA GAA ACA GAT CAC	336
Val Ser Thr Lys Thr Ser Ala Ala Asp Leu Val Thr Glu Thr Asp His	
95 100 105 110	
CGA GTA GAA GAC TTA ATT GTT TCT GAG TTG CGA AAG CGG TTC CCT TCA	384
Arg Val Glu Asp Leu Ile Val Ser Glu Leu Arg Lys Arg Phe Pro Ser	
115 120 125	
CAC AGG TTC ATT GCA GAA GAG GCC ACA GCC TCC GGG GCC AAG TGT GTG	432
His Arg Phe Ile Ala Glu Glu Ala Thr Ala Ser Gly Ala Lys Cys Val	
130 135 140	
CTC ACC CAC AGC CCG ACC TGG ATC ATC GAC CCC ATC GAC GGC ACC TGC	480
Leu Thr His Ser Pro Thr Trp Ile Ile Asp Pro Ile Asp Gly Thr Cys	
145 150 155	
AAC TTT GTG CAC AGA TTC CCC ACT GTG GCA GTT AGC ATC GGA TTT GCT	528
Asn Phe Val His Arg Phe Pro Thr Val Ala Val Ser Ile Gly Phe Ala	
160 165 170	
GTT CAC CAG GAG CTG GAA TTC GGA GTG ATT CAC CAC TGC ACA GAG GAG	576
Val His Gln Glu Leu Glu Phe Gly Val Ile His His Cys Thr Glu Glu	
175 180 185 190	
CGG CTG TAC ACC GGC AGG AGG GGC CAG GGC GCC TTT TGC AAT GGC CAG	624
Arg Leu Tyr Thr Gly Arg Arg Gly Gln Gly Ala Phe Cys Asn Gly Gln	
195 200 205	
AGG CTC CAG GTC TCC AGG GAG ACA GAT CTC GCA AAG GCC TTG GTT CTG	672
Arg Leu Gln Val Ser Arg Glu Thr Asp Leu Ala Lys Ala Leu Val Leu	



210	215	220	
ACA GAA ATC GGG CCC AAA CGT GAC	CCC GAT ACT CTG AAA GTA TTC CTG	720	
Thr Glu Ile Gly Pro Lys Arg Asp	Pro Asp Thr Leu Lys Val Phe Leu		
225	230	235	
AGC AAC ATG GAG CGG CTG CTG CAC	GCC AAG GCT CAT GGG GTC CGA GTG	768	
Ser Asn Met Glu Arg Leu Leu His	Ala Lys Ala His Gly Val Arg Val		
240	245	250	
ATT GGC AGC TCC ACC TTG GCG CTC	TGC TAC TTG GCC TCG GGG GCT GCT	816	
Ile Gly Ser Ser Thr Leu Ala Leu	Cys Tyr Leu Ala Ser Gly Ala Ala		
255	260	265	
GAT GCC TAT TAC CAG TTC GGC CTC	CAC TGC TGG GAT CTG GCA GCT GCC	864	
Asp Ala Tyr Tyr Gln Phe Gly Leu	His Cys Trp Asp Leu Ala Ala Ala		
275	280	285	
ACA GTC ATC ATC AGA GAA GCA GGT	GGC ATT GTG ATT GAC ACC TCA GGT	912	
Thr Val Ile Ile Arg Glu Ala Gly	Gly Ile Val Ile Asp Thr Ser Gly		
290	295	300	
GGA CCC CTT GAC CTC ATG TCG TGC	AGA GTG GTG GCT GCT GGC ACC AGA	960	
Gly Pro Leu Asp Leu Met Ser Cys	Arg Val Val Ala Ala Gly Thr Arg		
305	310	315	
GAG ATG GCA GTG CTC ATA GCT CAG	GCC CTA CAA ACC ATT AAC TAC GGC	1008	
Glu Met Ala Val Leu Ile Ala Gln	Ala Leu Gln Thr Ile Asn Tyr Gly		
320	325	330	
CGG GAC GAT GAG AAG TGAGCCGTAC	AGAGCTCTAA GGCTGACATG AGCAGCTCCC	1063	
Arg Asp Asp Glu Lys			
335			
TGGGAAAGAG CTGTCCGGGG GCTTGAGTTC	CGGGATAGTC TACCATAGCT GTCCCCGGAC	1123	
CTCGGTGCTT AGCTGATCCT CTCTAATCTC	GGGTAGCCCC TTTCCAGGTC GGTACATGGT	1183	
CTTTCATCAG AGCCAAACCC AAATCTTG TG	AGGTGTGTTA GTCACCCATC CTGGTTGTTC	1243	
GGAATGCAAA TCTCAGGTAA TAAAGCTTTA	GAACGAGCTC TCAGGCCCTC CCCTGCCCGT	1303	
GGTGATACAG AATGCAATAA ATCAGAACTT	TAAAAA	1339	

## (2) INFORMATION FOR SEQ ID NO:54:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Gly Ala Leu Glu Asp Ala Arg Leu Gly Gly Lys Ala Arg Gly Pro Cys  
 1 5 10 15  
 Gly Ala Ser Gly Ala Gly Gly Gly His Gln Phe Gly Ser Arg Glu Pro  
 20 25 30  
 Ala Ala Ala Arg Gly Thr Glu Gln Arg Val Ala Gly Pro Gly Glu Thr

35                      40                      45  
 Thr Met Lys Pro Asn Ser Glu Glu Glu Glu Leu Val Gln Gly Val  
 50                      55                      60  
 Gly Pro Trp Asp Glu Cys Phe Glu Val Ala Val Gln Leu Ala Leu Arg  
 65                      70                      75                      80  
 Ala Gly Gln Ile Ile Arg Lys Ala Leu Thr Glu Glu Lys His Val Ser  
 85                      90                      95  
 Thr Lys Thr Ser Ala Ala Asp Leu Val Thr Glu Thr Asp His Arg Val  
 100                      105                      110  
 Glu Asp Leu Ile Val Ser Glu Leu Arg Lys Arg Phe Pro Ser His Arg  
 115                      120                      125  
 Phe Ile Ala Glu Glu Ala Thr Ala Ser Gly Ala Lys Cys Val Leu Thr  
 130                      135                      140  
 His Ser Pro Thr Trp Ile Ile Asp Pro Ile Asp Gly Thr Cys Asn Phe  
 145                      150                      155                      160  
 Val His Arg Phe Pro Thr Val Ala Val Ser Ile Gly Phe Ala Val His  
 165                      170                      175  
 Gln Glu Leu Glu Phe Gly Val Ile His His Cys Thr Glu Glu Arg Leu  
 180                      185                      190  
 Tyr Thr Gly Arg Arg Gly Gln Gly Ala Phe Cys Asn Gly Gln Arg Leu  
 195                      200                      205  
 Gln Val Ser Arg Glu Thr Asp Leu Ala Lys Ala Leu Val Leu Thr Glu  
 210                      215                      220  
 Ile Gly Pro Lys Arg Asp Pro Asp Thr Leu Lys Val Phe Leu Ser Asn  
 225                      230                      235                      240  
 Met Glu Arg Leu Leu His Ala Lys Ala His Gly Val Arg Val Ile Gly  
 245                      250                      255  
 Ser Ser Thr Leu Ala Leu Cys Tyr Leu Ala Ser Gly Ala Ala Asp Ala  
 260                      265                      270  
 Tyr Tyr Gln Phe Gly Leu His Cys Trp Asp Leu Ala Ala Ala Thr Val  
 275                      280                      285  
 Ile Ile Arg Glu Ala Gly Gly Ile Val Ile Asp Thr Ser Gly Gly Pro  
 290                      295                      300  
 Leu Asp Leu Met Ser Cys Arg Val Val Ala Ala Gly Thr Arg Glu Met  
 305                      310                      315                      320  
 Ala Val Leu Ile Ala Gln Ala Leu Gln Thr Ile Asn Tyr Gly Arg Asp  
 325                      330                      335  
 Asp Glu Lys

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2285 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW059

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 9..1271

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1991..2162  
 (D) OTHER INFORMATION: /label= SAC\_24457

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

GTGACAG CCC GCC ACT CCA GGC GCG ATG CTG TTC TGG ACT GTG CTC AGC	50
Pro Ala Thr Pro Gly Ala Met Leu Phe Trp Thr Val Leu Ser	
1 5 10	
ATG GCT TTG AGT CTG CGG TTG GCA CTG GCG CAG AGC GGC ATA GAG CGC	98
Met Ala Leu Ser Leu Arg Leu Ala Leu Ala Gln Ser Gly Ile Glu Arg	
15 20 25 30	
GGT CCC ACA GCA TCA GCC CCC CAG GGG GAC CTG TTG TTC CTG CTG GAC	146
Gly Pro Thr Ala Ser Ala Pro Gln Gly Asp Leu Leu Phe Leu Leu Asp	
35 40 45	
AGC TCA GCC AGT GTG TCA CAC TAT GAG TTC TCA AGA GTT CGA GAG TTT	194
Ser Ser Ala Ser Val Ser His Tyr Glu Phe Ser Arg Val Arg Glu Phe	
50 55 60	
GTG GGG CAG CTG GTG GCT ACG ATG CCT TTC GGA CCC GGG GCT CTG CGT	242
Val Gly Gln Leu Val Ala Thr Met Pro Phe Gly Pro Gly Ala Leu Arg	
65 70 75	
GCT AGT CTG GTG CAC GTG GGC AGC CGG CCT CAC ACG GAG TTT ACT TTT	290
Ala Ser Leu Val His Val Gly Ser Arg Pro His Thr Glu Phe Thr Phe	
80 85 90	
GAT CAG TAC AGT TCA GGC CAG GCC ATA CAG GAT GCT GTA CGT GTT GCG	338
Asp Gln Tyr Ser Ser Gly Gln Ala Ile Gln Asp Ala Val Arg Val Ala	
95 100 105 110	
CCC CAA CGT ATG GGT GAC ACC AAC ACG GGC CTG GCA CTG GCT TAT GCC	386
Pro Gln Arg Met Gly Asp Thr Asn Thr Gly Leu Ala Leu Ala Tyr Ala	
115 120 125	
AAA GAA CAA TTA TTT GCT GAG GAA GCA GGT GCC CGG CTA GGG GTT CCC	434
Lys Glu Gln Leu Phe Ala Glu Glu Ala Gly Ala Arg Leu Gly Val Pro	
130 135 140	
AAG GTA CTG GTG TGG GTG ACA GAT GGG GCC TCC AGT GAC TCT GTG GGC	482
Lys Val Leu Val Trp Val Thr Asp Gly Ala Ser Ser Asp Ser Val Gly	
145 150 155	

CCC Pro 160	CCT Pro 160	ATG Met	CAG Gln	GAG Glu	CTC Leu	AAG Lys	GAC Asp	CTG Leu	GGT Gly	GTC Val	ACC Thr	ATC Ile	TTC Phe	ATT Ile	GTC Val	530
AGC Ser 175	ACT Thr	GGC Gly	CGA Arg	GGC Gly	AAC Asn	CTG Leu	TTG Leu	GAG Glu	CTG Leu	TTG Leu	GCA Ala	GCT Ala	GCC Ala	TCA Ser	GCT Ala	578
CCG Pro	GCT Ala	GAG Glu	AAG Lys	CAC His	CTA Leu	CAC His	TTT Phe	GTG Val	GAT Asp	GTG Val	GAT Asp	GAC Asp	CTT Leu	CCT Pro	ATC Ile	626
ATT Ile	GCC Ala	CGG Arg	GAG Glu	CTT Leu	CGG Arg	GGT Gly	GCC Ala	ATT Ile	ATT Ile	GAT Asp	GCG Ala	ATG Met	CAG Gln	CCA Pro	CAT His	674
CAG Gln	CTT Leu	CAT His	GCT Ala	TCG Ser	GAG Glu	ATT Ile	CTG Leu	TCC Ser	AAT Asn	GGC Gly	TTC Phe	CGC Arg	CTG Leu	TCC Ser	TGG Trp	722
CCG Pro 240	CCC Pro	CTG Leu	CTG Leu	ACA Thr	GCG Ala	GAC Asp	TCT Ser	GGT Gly	TAC Tyr	TAC Tyr	GTG Val	CTG Leu	GAG Glu	TTG Leu	GTG Val	770
CCC Pro 255	AGT Ser	GGC Gly	AAA Lys	CTG Leu	GCA Ala	GCC Ala	ACA Thr	AGA Arg	CGC Arg	CAA Gln	CAG Gln	CTG Leu	CCC Pro	GGG Gly	AAT Asn	818
GCT Ala	ACC Thr	AGC Ser	TGG Trp	ACC Thr	TGG Trp	ACC Thr	GAC Asp	CTC Leu	AAC Asn	CCA Pro	GAC Asp	ACA Thr	GAT Asp	TAC Tyr	GAA Glu	866
GTA Val	TCG Ser	CTG Leu	TTG Leu	CCG Pro	GAG Glu	TCC Ser	AAT Asn	GTG Val	CGC Arg	CTC Leu	CTG Leu	AGG Arg	CCC Pro	CAG Gln	CAC His	914
TTG Leu	CGA Arg	GTA Val	CGC Arg	ACA Thr	CTG Leu	CAA Gln	GAG Glu	GAG Glu	GCA Ala	GGG Gly	CCA Pro	GAA Glu	CGC Arg	ATC Ile	GTC Val	962
ATC Ile 320	TCG Ser	CAT His	ACT Thr	AGG Arg	CCG Pro	CGC Arg	AGC Ser	CTC Leu	CGT Arg	GTA Val	AGT Ser	TGG Trp	GCC Ala	CCC Pro	GCA Ala	1010
CTT Leu 335	GGC Gly	CCG Pro	GAC Asp	TCC Ser	ACT Thr	CTC Leu	GGC Gly	TAC Tyr	CTT Leu	GTA Val	CAG Gln	CTC Leu	GGA Gly	CCT Pro	CTG Leu	1058
CAG Gln	GGC Gly	GGA Gly	TCC Ser	CTA Leu	GAA Glu	CAC His	GTG Val	GAG Glu	GTG Val	CCA Pro	GCT Ala	GGC Gly	CAG Gln	AAC Asn	AGC Ser	1106
ACT Thr	ACC Thr	ATC Ile	CAG Gln	GGC Gly	CTG Leu	ACG Thr	CCC Pro	TGC Cys	ACC Thr	ACT Thr	TAC Tyr	CTG Leu	GTG Val	ACT Thr	GTG Val	1154
ACT Thr	GCC Ala	GCC Ala	TTC Phe	CGC Arg	TCG Ser	GGC Gly	CGC Arg	CAG Gln	AGG Arg	GCG Ala	CTG Leu	TCG Ser	GCT Ala	AAG Lys	GCC Ala	1202
TGT Cys	ACG Thr	GCA Ala	TCT Ser	GGC Gly	GAG Glu	CGG Arg	ATC Ile	CGT Arg	GTC Val	CCG Pro	CAG Gln	GCC Ala	ATG Met	CGG Arg	CCG Pro	1250

400	405	410	
GAG GCC GGA CTG CGG GAG CCC TGACCTTCCT GCCTCGTCCA CCCGAGGGGC			1301
Glu Ala Gly Leu Arg Glu Pro			
415	420		
CCTCTTCCT AATCCAGTGA GAGAGGCACC GCTGCTCGTG GTTTTCTTG TGGATGGAGT			1361
GGGGTGGGGA GATGGGGGTG CTGGTCCTAC CTTTGACCTG CGTAACCTCT CCGGTCGTTT			1421
CCCCACTGGT CATCACCGCC CTGCGCTGAC TCCCCGGAAA CCTGTAGCAA CCCTGGTAGC			1481
CTCACGCGCA ATGACAATCC TCTCCGGTTG CCAGTGAGT TGAGCACACG GTGGTCCTTG			1541
GGCAACATTT GGCAGGGGGA TGGACAGTGT TTGAGGTCAG GTTGAGACCC AGGAGAAGCA			1601
TTCAGGAGAG GAGGCCACAG AGTTTTCTAC CTGTGCCAAA GACTGGGCCC TCTGGTGGCA			1661
AGGACTACAC ATGGCTTGGA GGAAATGTCC AGGACCCTTC AAGTCCTGCC TGTGCCTAGA			1721
AAGTGGGTAG GAGAAAGGGA GGAGAGAGTA GTGTAGGCAA GGTTCCTTCA GACTTCCTTA			1781
GGGCAAGGAA AGGTAGGGAG AGATACCGGG AGGCTGATGA TGTGCCAAC TAGTTTTCAT			1841
CAAGATTTTC TGCCAGCCTG GAGGCCAGGA TCTGTCAGGG TCACTGACTC TGCCTTCCTG			1901
CCCAGGACCT GCACTGGGCC CCGCTATCAG TCGGGGGGGG GGGGTGCAGA GTCTTCACAG			1961
GAATGGGGGA TGAGACCTTG GCATGTAGTA CATTGGGGAT AGGAGAGCCC TGCCGTGACA			2021
GACTTACAGG GAGTCTCCTG CTTTAGTGTA GGGAGCAAGG TGACATGCAG GTGGGCTACC			2081
TCCTGTCATC ACTACTGCCC TGGGGCATCT GACAGATACC TAAGGGTAGT CAGGAACAGG			2141
CTTCTCTCC AGTCCCTATG TACGCAGAGC CCCTCCTCCC CAGAACCTCT TGCCCTAACC			2201
TAAGCTTACT CCATCTCTCT TCCCCACTAA TGACCCAGAC TCTAACAATA ATACAGTAAG			2261
CCAGATGTAA ACTGTGAAGT CGAC			2285

## (2) INFORMATION FOR SEQ ID NO:56:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 421 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Pro	Ala	Thr	Pro	Gly	Ala	Met	Leu	Phe	Trp	Thr	Val	Leu	Ser	Met	Ala
1				5					10					15	
Leu	Ser	Leu	Arg	Leu	Ala	Leu	Ala	Gln	Ser	Gly	Ile	Glu	Arg	Gly	Pro
		20						25					30		
Thr	Ala	Ser	Ala	Pro	Gln	Gly	Asp	Leu	Leu	Phe	Leu	Leu	Asp	Ser	Ser
	35					40						45			
Ala	Ser	Val	Ser	His	Tyr	Glu	Phe	Ser	Arg	Val	Arg	Glu	Phe	Val	Gly
50						55						60			

Gln Leu Val Ala Thr Met Pro Phe Gly Pro Gly Ala Leu Arg Ala Ser  
 65 70 75 80  
 Leu Val His Val Gly Ser Arg Pro His Thr Glu Phe Thr Phe Asp Gln  
 85 90 95  
 Tyr Ser Ser Gly Gln Ala Ile Gln Asp Ala Val Arg Val Ala Pro Gln  
 100 105 110  
 Arg Met Gly Asp Thr Asn Thr Gly Leu Ala Leu Ala Tyr Ala Lys Glu  
 115 120 125  
 Gln Leu Phe Ala Glu Glu Ala Gly Ala Arg Leu Gly Val Pro Lys Val  
 130 135 140  
 Leu Val Trp Val Thr Asp Gly Ala Ser Ser Asp Ser Val Gly Pro Pro  
 145 150 155 160  
 Met Gln Glu Leu Lys Asp Leu Gly Val Thr Ile Phe Ile Val Ser Thr  
 165 170 175  
 Gly Arg Gly Asn Leu Leu Glu Leu Leu Ala Ala Ala Ser Ala Pro Ala  
 180 185 190  
 Glu Lys His Leu His Phe Val Asp Val Asp Asp Leu Pro Ile Ile Ala  
 195 200 205  
 Arg Glu Leu Arg Gly Ala Ile Ile Asp Ala Met Gln Pro His Gln Leu  
 210 215 220  
 His Ala Ser Glu Ile Leu Ser Asn Gly Phe Arg Leu Ser Trp Pro Pro  
 225 230 235 240  
 Leu Leu Thr Ala Asp Ser Gly Tyr Tyr Val Leu Glu Leu Val Pro Ser  
 245 250 255  
 Gly Lys Leu Ala Ala Thr Arg Arg Gln Gln Leu Pro Gly Asn Ala Thr  
 260 265 270  
 Ser Trp Thr Trp Thr Asp Leu Asn Pro Asp Thr Asp Tyr Glu Val Ser  
 275 280 285  
 Leu Leu Pro Glu Ser Asn Val Arg Leu Leu Arg Pro Gln His Leu Arg  
 290 295 300  
 Val Arg Thr Leu Gln Glu Glu Ala Gly Pro Glu Arg Ile Val Ile Ser  
 305 310 315 320  
 His Thr Arg Pro Arg Ser Leu Arg Val Ser Trp Ala Pro Ala Leu Gly  
 325 330 335  
 Pro Asp Ser Thr Leu Gly Tyr Leu Val Gln Leu Gly Pro Leu Gln Gly  
 340 345 350  
 Gly Ser Leu Glu His Val Glu Val Pro Ala Gly Gln Asn Ser Thr Thr  
 355 360 365  
 Ile Gln Gly Leu Thr Pro Cys Thr Thr Tyr Leu Val Thr Val Thr Ala  
 370 375 380  
 Ala Phe Arg Ser Gly Arg Gln Arg Ala Leu Ser Ala Lys Ala Cys Thr  
 385 390 395 400

Ala Ser Gly Glu Arg Ile Arg Val Pro Gln Ala Met Arg Pro Glu Ala  
 405 410 415

Gly Leu Arg Glu Pro  
 420

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2521 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW061

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..1422

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (2089..2287)
- (D) OTHER INFORMATION: /label= SAC\_24029

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GTCGAC GCG ACG GTG GCG GAG CAG TGG GTG CTG GTG GAG ATG GTG CAG	48
Ala Thr Val Ala Glu Gln Trp Val Leu Val Glu Met Val Gln	
1 5 10	
GCG CTG TAC GAG GCT CCA GCA TAC CAT CTA ATT CTG GAA GGA ATC CTC	96
Ala Leu Tyr Glu Ala Pro Ala Tyr His Leu Ile Leu Glu Gly Ile Leu	
15 20 25 30	
ATA CTT TGG ATA ATC AGA CTC GTT TTC TCT AAA ACT TAC AAG TTG CAG	144
Ile Leu Trp Ile Ile Arg Leu Val Phe Ser Lys Thr Tyr Lys Leu Gln	
35 40 45	
GAG CGT TCT GAC CTT ACA GCC AAG GAA AAG GAA GAA CTG ATT GAA GAA	192
Glu Arg Ser Asp Leu Thr Ala Lys Glu Lys Glu Glu Leu Ile Glu Glu	
50 55 60	
TGG CAG CCA GAG CCC CTC GTC CCC CCG GTC TCC AGG AAC CAT CCT GCT	240
Trp Gln Pro Glu Pro Leu Val Pro Pro Val Ser Arg Asn His Pro Ala	
65 70 75	
CTC AAC TAC AAC ATC GTC TCC GGC CCT CCA ACC CAT AAC ATC GTG GTG	288
Leu Asn Tyr Asn Ile Val Ser Gly Pro Pro Thr His Asn Ile Val Val	
80 85 90	
AAC GGA AAA GAG TGT GTC AAC TTT GCC TCC TTT AAT TTC CTC GGG TTG	336
Asn Gly Lys Glu Cys Val Asn Phe Ala Ser Phe Asn Phe Leu Gly Leu	
95 100 105 110	

CTG GCC AAC CCT CGG GTT AAG GCT GCG GCT TTT GCA TCA TTA AAG AAG Leu Ala Asn Pro Arg Val Lys Ala Ala Phe Ala Ser Leu Lys Lys	384
115 120 125	
TAT GGC GTG GGT ACC TGT GGT CCC AGA GGG TTT TAT GGC ACA TTT GAT Tyr Gly Val Gly Thr Cys Gly Pro Arg Gly Phe Tyr Gly Thr Phe Asp	432
130 135 140	
GTC CAT CTG GAT TTG GAA GAG CGC CTG GCA AAG TTT ATG AAG ACC GAA Val His Leu Asp Leu Glu Glu Arg Leu Ala Lys Phe Met Lys Thr Glu	480
145 150 155	
GAA GCT ATC ATT TAC TCC TAT GGC TTC TCC ACC ATA GCC AGT GCG ATT Glu Ala Ile Ile Tyr Ser Tyr Gly Phe Ser Thr Ile Ala Ser Ala Ile	528
160 165 170	
CCT GCG TAC TCT AAG AGA GGG GAC ATT GTT TTT GTG GAC AGT GCA GCC Pro Ala Tyr Ser Lys Arg Gly Asp Ile Val Phe Val Asp Ser Ala Ala	576
175 180 185 190	
TGC TTT GCT ATC CAG AAA GGA CTA CAG GCA TCA CGC AGT GAC ATT AAG Cys Phe Ala Ile Gln Lys Gly Leu Gln Ala Ser Arg Ser Asp Ile Lys	624
195 200 205	
TTA TTC AAG CAC AAT GAC GTA GCT GAC CTG GAG CGA CTG CTA AAA GAA Leu Phe Lys His Asn Asp Val Ala Asp Leu Glu Arg Leu Leu Lys Glu	672
210 215 220	
CAA GAG ATT GAA GAT CAA AAG AAT CCT CGA AAG GCC CGT GTG ACT CGG Gln Glu Ile Glu Asp Gln Lys Asn Pro Arg Lys Ala Arg Val Thr Arg	720
225 230 235	
CGA TTC ATC GTC GCG GAA GGA TTG TAT ATG AAC ACT GGA ACC ATC TGC Arg Phe Ile Val Ala Glu Gly Leu Tyr Met Asn Thr Gly Thr Ile Cys	768
240 245 250	
CCT CTT CCA GAA CTG GTG AGG TTA AAG TAT AAA TAC AAA GCA AGG ATC Pro Leu Pro Glu Leu Val Arg Leu Lys Tyr Lys Tyr Lys Ala Arg Ile	816
255 260 265 270	
TTC CTG GAG GAG AGC CTG TCG TTC GGA GTC CTT GGG GAG CAT GGG CGA Phe Leu Glu Glu Ser Leu Ser Phe Gly Val Leu Gly Glu His Gly Arg	864
275 280 285	
GGA GTC ACC GAG CAC TAT GGG ATC AGT ATT GAC GAC ATC GAC CTT ATC Gly Val Thr Glu His Tyr Gly Ile Ser Ile Asp Asp Ile Asp Leu Ile	912
290 295 300	
AGT GCT AAC ATG GAG AAT GCG CTC GCT TCT GTT GGG GGC TTC TGC TGT Ser Ala Asn Met Glu Asn Ala Leu Ala Ser Val Gly Gly Phe Cys Cys	960
305 310 315	
GGC CGC TCT TTC GTG GTT GAC CAT CAG CGG CTC TCT GGC CAA GGA TAC Gly Arg Ser Phe Val Val Asp His Gln Arg Leu Ser Gly Gln Gly Tyr	1008
320 325 330	
TGC TTT TCA GCT TCA CTG CCC CCC CTG CTA GCC GCT GCC GCC ATT GAG Cys Phe Ser Ala Ser Leu Pro Pro Leu Leu Ala Ala Ala Ala Ile Glu	1056
335 340 345 350	
GCC CTC AAC ATC ATG GAG GAG AAC CCA GGG ATT TTT GCA GTT TTA AAG Ala Leu Asn Ile Met Glu Glu Asn Pro Gly Ile Phe Ala Val Leu Lys	1104
355 360 365	



AAA AAG TGC CAG ACC ATC CAC AAG TCC CTA CAA GGG GTT TCC GGT TTA	1152
Lys Lys Cys Gln Thr Ile His Lys Ser Leu Gln Gly Val Ser Gly Leu	
370 375 380	
AAA GTG GTG GGG GAG TCC CTT TGC CCA GCG CTT CAC CTC CAG CTG GAA	1200
Lys Val Val Gly Glu Ser Leu Cys Pro Ala Leu His Leu Gln Leu Glu	
385 390 395	
GAG AGC ACG GGC TCT CGG GAG AGA GAT ATG AAG CTG CTT CAG GAG ATT	1248
Glu Ser Thr Gly Ser Arg Glu Arg Asp Met Lys Leu Leu Gln Glu Ile	
400 405 410	
GTA GAG CAA TGC ATG AAT AAG GGC ATC GCA TTG ACT CAG GCA CGC TAC	1296
Val Glu Gln Cys Met Asn Lys Gly Ile Ala Leu Thr Gln Ala Arg Tyr	
415 420 425 430	
TTG GAC AAG GAA GAG AAG TGC CTT CCT CCT CCA AGC ATC AGG GTT GTG	1344
Leu Asp Lys Glu Glu Lys Cys Leu Pro Pro Ser Ile Arg Val Val	
435 440 445	
GTC ACC GTG GAG CAG ACA GAC GAA GAG CTG CAG AGG GCT GCA GCC ACC	1392
Val Thr Val Glu Gln Thr Asp Glu Glu Leu Gln Arg Ala Ala Thr	
450 455 460	
ATC AGG GAG GCG GCC CAG GCT GTG CTT TTG TAGGCTCCTG CCCAGTGCTT	1442
Ile Arg Glu Ala Ala Gln Ala Val Leu Leu	
465 470	
GCGACCATGT CACCTGCAGA CAGAACTACT CAGACCTCCA GCTGCCCAGG GCGTGGAGTG	1502
CTGCCCAGGG CGTGGCTGCC CTTGAGTCCA CAGCCAGAAT GGATGGCTTT ACCCAGCAGG	1562
CTTCAGAGG ATGGCAGACG ACGGTGTAAC ACTGGTGTGT ATGCGGTGCC ATGGCTCTGT	1622
CTCTTATACT CGCCTTTGTT TAAAACTGGA GCTTCCGTTT TCCCTCTCCC GCCTGTGGGA	1682
ACAATTCGAT AAGCACTCCT CTGTATTCCA TACAAAAGGA CTTGCAGCTG AGAGGAAGTC	1742
TCCACGCACC ACAGTTACTG TGGGATGATA GACTGTGGCC TCCCAGGTCC TTCGGGTAGA	1802
GCTCAGTGGT GGCTGCTGTG CAGACAATC CTGTGTGTCA AAAGTGCAT AGAACATGTT	1862
GATCCTAGAA CTCTGTGTGC TCCAGTTACT GATGCTGTGC TGTCTAATCT GAGCTTCACT	1922
AACAATAGTC ATACCGCTCC CTAAGCAATA ATCTGAATCA AACATTATTT TATTTTGTG	1982
TAATTGACTT TATATTTTTT ATATATTCTC TAGTCAGTTT TGGGAAGAGG GTTATTATTC	2042
TGGTCCCCAA ATATATAAAT CATATCCTTG GTATTTTTTA ATATGTACGG TGTCCTTTTT	2102
CATGGTGGTT TTGTCCAGGA ACATAGAACT GTGGGCCTTT TCTAAGCTGA GAAAAGAGAG	2162
AACTTTCTTC TTTCGTTCTA TAAATTCCAG ATGCCTCAGA AAATAGGAAT GCTCTGAAAA	2222
CTGCTGCTGT CCTTAGCGAG CGTGCTGAGG GGTGTAACGG CTTGTGACAT GTGCCGCGAT	2282
CGGTACATG GAAGGCCATT ACTCCCTCTC GCTGCACAGG CAGAGGAATG GGGCCTCTAA	2342
CCGTCGTTTT CGGATGTTCA TTCATGTTAA CAGTGGACGG CATCTTCCTA AGTGATCTGT	2402
GACATGTTTA ATGTGTCTAT TTCATACGTG CCTGTGGAAT GCTGCTGCTG TGAGGGTACG	2462

GGGATGAGTC TGTTTTGGTT TGGTTTTTAA TGAAATAAAC CTCCAAGAGC CTACAAAAA 2521

## (2) INFORMATION FOR SEQ ID NO:58:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Ala Thr Val Ala Glu Gln Trp Val Leu Val Glu Met Val Gln Ala Leu  
 1 5 10 15  
 Tyr Glu Ala Pro Ala Tyr His Leu Ile Leu Glu Gly Ile Leu Ile Leu  
 20 25 30  
 Trp Ile Ile Arg Leu Val Phe Ser Lys Thr Tyr Lys Leu Gln Glu Arg  
 35 40 45  
 Ser Asp Leu Thr Ala Lys Glu Lys Glu Glu Leu Ile Glu Glu Trp Gln  
 50 55 60  
 Pro Glu Pro Leu Val Pro Pro Val Ser Arg Asn His Pro Ala Leu Asn  
 65 70 75 80  
 Tyr Asn Ile Val Ser Gly Pro Pro Thr His Asn Ile Val Val Asn Gly  
 85 90 95  
 Lys Glu Cys Val Asn Phe Ala Ser Phe Asn Phe Leu Gly Leu Leu Ala  
 100 105 110  
 Asn Pro Arg Val Lys Ala Ala Ala Phe Ala Ser Leu Lys Lys Tyr Gly  
 115 120 125  
 Val Gly Thr Cys Gly Pro Arg Gly Phe Tyr Gly Thr Phe Asp Val His  
 130 135 140  
 Leu Asp Leu Glu Glu Arg Leu Ala Lys Phe Met Lys Thr Glu Glu Ala  
 145 150 155 160  
 Ile Ile Tyr Ser Tyr Gly Phe Ser Thr Ile Ala Ser Ala Ile Pro Ala  
 165 170 175  
 Tyr Ser Lys Arg Gly Asp Ile Val Phe Val Asp Ser Ala Ala Cys Phe  
 180 185 190  
 Ala Ile Gln Lys Gly Leu Gln Ala Ser Arg Ser Asp Ile Lys Leu Phe  
 195 200 205  
 Lys His Asn Asp Val Ala Asp Leu Glu Arg Leu Leu Lys Glu Gln Glu  
 210 215 220  
 Ile Glu Asp Gln Lys Asn Pro Arg Lys Ala Arg Val Thr Arg Arg Phe  
 225 230 235 240  
 Ile Val Ala Glu Gly Leu Tyr Met Asn Thr Gly Thr Ile Cys Pro Leu  
 245 250 255  
 Pro Glu Leu Val Arg Leu Lys Tyr Lys Tyr Lys Ala Arg Ile Phe Leu

260	265	270
Glu Glu Ser Leu Ser Phe Gly Val	Leu Gly Glu His Gly Arg Gly Val	
275	280	285
Thr Glu His Tyr Gly Ile Ser Ile Asp Asp Ile Asp Leu Ile Ser Ala		
290	295	300
Asn Met Glu Asn Ala Leu Ala Ser Val Gly Gly Phe Cys Cys Gly Arg		
305	310	315
Ser Phe Val Val Asp His Gln Arg Leu Ser Gly Gln Gly Tyr Cys Phe		
325	330	335
Ser Ala Ser Leu Pro Pro Leu Leu Ala Ala Ala Ala Ile Glu Ala Leu		
340	345	350
Asn Ile Met Glu Glu Asn Pro Gly Ile Phe Ala Val Leu Lys Lys Lys		
355	360	365
Cys Gln Thr Ile His Lys Ser Leu Gln Gly Val Ser Gly Leu Lys Val		
370	375	380
Val Gly Glu Ser Leu Cys Pro Ala Leu His Leu Gln Leu Glu Glu Ser		
385	390	395
Thr Gly Ser Arg Glu Arg Asp Met Lys Leu Leu Gln Glu Ile Val Glu		
405	410	415
Gln Cys Met Asn Lys Gly Ile Ala Leu Thr Gln Ala Arg Tyr Leu Asp		
420	425	430
Lys Glu Glu Lys Cys Leu Pro Pro Pro Ser Ile Arg Val Val Val Thr		
435	440	445
Val Glu Gln Thr Asp Glu Glu Leu Gln Arg Ala Ala Ala Thr Ile Arg		
450	455	460
Glu Ala Ala Gln Ala Val Leu Leu		
465	470	

## (2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 598 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW062
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: complement (87..223)
  - (D) OTHER INFORMATION: /label= SAC\_24443

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```

TTTTTGTGAG CAAATCCAAA TTTATTTTAA TGCCATGTCA TTTCAAAAAC CCCAGCCTTG      60
GTTCCCTGGA ACAGCATGCC AGAGGTACAA AGTGTAACCT TCCTCTATAA ACCCCCAGCA      120
AGTCATCCAA GTCCTCAGCT TCAGAAAGTC AGTTCTCTTC AGTCTACTCC TCTGGTTCCT      180
GGTTTTTCTT TCAAGGAAGG CAACTTAAAA ATTAGGTAGC TGAGGTTCCA ATGTTGGTGG      240
CTAACCTTGC CTCCTCTTTT GTTCTGTAGT AGGGCCACAA ACCTTGACCA GGCAGAGTAG      300
TAGAAAAGTA GAAAGAGGGG CTTGACGACG GTGGATTTTG ACTCCTGATT TTATTATTCA      360
ATTTCTTTTT CTACTAAAAG TAGTCTTCGG TGGTTGGGAA GCCTGGCCTC CCAACACCAG      420
AGTCAGTCGG AGCTGGTTTT TTTGTTGAAA GGAGTGGGCG GGTGGGTGGG GGACCGGGAT      480
GAGGGCAGAA CCCCGCTCTG CTGGTAGTCT TGGGTGGAGA AGACGAACTG CACTTGACAG      540
AGCCTGGGGG TGCGGTGGGA GGGGGTGAGG CAGGAGTGAC AGCTGGGGAG GGGACCCA      598

```

## (2) INFORMATION FOR SEQ ID NO:60:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW066

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 759..906
- (D) OTHER INFORMATION: /label= SAC\_24521

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```

ATTCGCGGCC GGCTCGACTT TTTTTTTTGT GATATAAAAA ATTTATTTTA GAAATCACAT      60
CTTTAAAAAA CAACATAGAA GGTATCCTG ACCAGATATA AATTCTTCAT TCAATGCCAG      120
GGTGTATATG GGTACAATTA GCGGGATGTT CTGTGAGGAG AATGAGTCTG TGTGATACTC      180
TCGACATGTA CACCAGCTCA AACGTTAATG CGGCTTTTAA AAGGAAGGAG AGAGGAAGGC      240
AGCGAGAGAG GAGGAAACTG TTTTGGAAC AGAAGAGCAT CTCCGTCCTG CACCTGTCAG      300
CCCTGAAACT CCCGGCCCTT GAGACCAGCC CAGCAGAGGC TGCACGAANN NGCATCATTT      360
TAATTCACAA GTGACTGCAG CACCCTGGCA TCTTGTGCTG GGAGTGTTGG TGGGAACAGA      420

```

GGTGACAAGT CTGCTCTGTA GTGTCATAGC CAGAGAGCAA CAGGGACTCT AGAGGAAGAA 480  
GAGCAAAGCT TGCCTAGACA GAACGGGTG TCCCTGTCCC TGTCCATCAG CTTTCTTATA 540  
ACCTATGGNT AGCCAGGATC AAATACCTGA AAGTCAGAAT GTAAAATAAA AGTGTCAGTT 600  
AACACTTATT GCGGACAGAA ATACAGCTCT TTACGCCTAC AACCATAAAC CCAGTGGTGA 660  
GGCTGGGACA GACCTTTTAT GCCGAAGTTC TCACAACTCG GGAAGTACTT ACTGGGGCCT 720  
GGTCACTGCT TAATTGGTGA CCTCACACCA TTTCAGGTAC TTCTGCCCAG CTCCTGATTC 780  
CTCTTCAGGA GAAACAGGAA GCAGACTTTG TTAGTAAGTG GGTATGCAG GGCTGGGCCG 840  
CCTACATGGG ACTTAACACA GAGAGAATAC TGCCAGCTTC CCAAACACTT CAAGAGAGAA 900  
AAATGTCCAG ACGTGGGTGC TGCAGCAACC TTCCAAGAGT GAGGCCCCCTG CACCCTCCTG 960  
CACCCTCGGC TGGGGACAGT GGGCTCTGGT GGCTTGAAGT GAGTAGATCG GATCATTTCC 1020  
TCCTCTTTCT GAGACAAGGC TTCCTGTGT AGCCCNNGCT GCCCTGGAGC TCACTATTTA 1080  
GACCAGGCTA GGATCTCACA GAGAACCACC TGCCTCTGGC CTCTCAAGAG CTGGGATTAA 1140  
GGTGTGAAAA GGGGGCTAAT GTAGCCCAGG CTGGCTCCTG CTCCACAGGG CTTGCAGGCC 1200  
AGTTCTGGAT CCTACCAGCA AGTACAAGCT GGTACCCAGA GGACGCTGCA GGATGGAAGG 1260  
ACGGAAGGCC AGGAGCCTTC GTTTGATTTG ACGGGAAGGA GGAGGAACAG TGAGGTTGTG 1320  
CATCTGAGCT TTATTACCAG GACGATGATG TCACCTCAGG AAGGGACTGG GATCCCTGCA 1380  
AGTGGTTCCA CATGGGACCC AGGTCCTCAG AGATGTGGCC CTCCTTGTGT GCCAAAATGC 1440  
ATCTGGAAAG AAAATACAGA AGGGGAGGGA CAGCATAGGA GGAGACAAGT GTGACCCGAG 1500  
GCCCAGGAAG TACTTCCATA TTCTGTCCTT CTCCAGGCAC GAGAGTTGAA GGCCGTGCCA 1560  
AGCCCTGCAC TCGCTAAGCT GGTGAAGACC ACACCAAAT CACTGTGACA GGACAGCTGC 1620  
ACAGATCTGT TTCCGACTTT TCCCACCAG ATTCTACACC AAGGTAAAAA AGAGCTTCCA 1680  
CCACAAATCC AAGGCAGGAG CTGGGCGGCC ACACACCCGC CACAGATGGT GCAAGTTGCT 1740  
CACTTCCGCA GCCGGTGGCA GGACAGGGAG GAGGGCAGAG TCCTGAGACA AGAACCCAG 1800  
AGCGGGGCA GTTCAGTCAG TGGTGGACCT CCGGTACCAG AACCGAGCTC GGAAGTCGTC 1860  
GCTGCACGTC ATGAAGCCAG GGTGGTCCG GAGTGCAGT ATGCAGCGCA CAATGTTGTT 1920  
GTGGCCAGG GACAGCAGGT TCCTGCGCTC AGCTGTCTTA GAGTCCAGC AGCACAGGCT 1980  
GATGGTCCTT TCGTCCGGCA GCAAAATATA GTCCTCAGTG TGTTGAACA CAGCCTGTGT 2040  
CCGGGTGCAC TTGCCGTCCA CTCAGCCCTG CGCCTGTGTA CCTCACCAGT GTCCGTCCCG 2100  
TGGATATTTT CCCAGAGTTT AGCTACAGAG TCTTTYCCNN GAGTTTAGGC CTACAGAGTC 2160  
TTTTCCACTT GAGAGGATGT ACTTGGAAAT CTTGGAGAAA ATGGCAGAAC AGACTTCTGC 2220  
CCCGTCGTGA GCTTTCTCAA AGGTTGTGAT GCAGCGATTG GAGACGCCGT CCCACAGCTT 2280

GATGCAGCCG TCCTTGCTGC CAGTGACATA CATGTTGGCA CTAGGGTTAT AGTTGACAGA	2340
GCAGATGGCG TCGGTGTGCT GGTCTTGGGG ATTGCAGGAG ACGAAACACT GGAAGGTGTT	2400
GATGTCATAG AGGCGAAGCG TGGGATGCTG GGTCCCGACC AGGATGAAGT CTCCAGAAGG	2460
GTGGAACGAG ATGGAGCGCA GCATCTCAGC TTCCTGAATG TATTTAAAGG CTCTTTTTC	2520
AGAGGGTTTG GAATAATCAA ACAATTTTAG AGTATAATCC CTTGAGCCGG AGGCCAGGAT	2580
CTGTTCTGTT GGGTGAAAAG CGAGACACGT GACTTCATCC ACATGGTCGT AAAGAGTTCG	2640
GATCACTGGG TGGTTCTCCA TGTCTGTTG TGCAGTCTCA TTCATCATGA CCCTCGATGG	2700
GCATGGCGCT CTTGGCCAGC ATCCTCTCTG TGTCCAGTAT CTTGATGGAA GCATCCGCAG	2760
ACCCTGTGGC TATCAACTGC CCATCTCTGC TGTAGGTGGC CACTCGGCAT GGGCCTTGT	2820
GGGAGGTGAC ATAACAGGTT TCATACTCCG AAGCCTCGGG TGACATTGTT TGGACGTCTG	2880
CATCAAATC CAGGTCAATN CTGTGCCAGG GGCAACCGTG TCAGAGCGAC CAATGGCATA	2940
CTGAACTGCA GTGTCATCGT TCTCCATCCC TAGCTTGATG AGGTGCAGGA GCTGCTCCGA	3000
GGGCGCACAC ACCGACTGTG GCTTGATCTC GTTGATGAGG CCATTGCGAT GCTAATGTAT	3060
CCGTCATAGA GCAGCTGGCT GATGATCAGC TTGTAGAGCT GCTGGCGGTC CTTCAGGCCC	3120
ACTTTGGTTC TGTACATTCT GCACAAGAGG ATGACAGTTC CCCTGACTTC TCCGAGCTGA	3180
TCACCTCCGC GACAGCCTTG AATCCAAAGG ACTGTGAGTG ACCGGCAAAC GCCCGANCC	3240
TGGAGAGCAC NGCGGGCGAT AGAACGGCCG TAGATCGGAC AGGCCGTTGG TGCACCGTCG	3300
ACCGG	3305

## (2) INFORMATION FOR SEQ ID NO:61:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2099 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW069

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 1360..1893

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 1744..1885  
 (D) OTHER INFORMATION: /label= SAC\_24477

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GTCGACGCAT CCAGGCAGGC ATGGAGCAGG TAGGGCTCAG AGTTCTACAT CTTCATCTGG	60
AGGCTACTAT GAGGAGACCG ACTTCCAGGC AGCTAGGAAG CCCATGCTCA CAGTGACAGA	120
CCTTCTCCAA CAAGGTTATA CCTCCTAATC GTGCCGCTCC CTGGGCCAAG CATATGCAAA	180
CCTTCACACA GGTCTTCCTC CTACACAGTG GGGCTTCTGG GCCCCTCA CGGCCAGTAC	240
ATGGCAAAAG ACATTAAAGA TGCACCTGGA CCAGCTCAGG GCCCTGTTCT GCAGCCAGGT	300
ATCATGGATG GAGGAAGTCC TGA CTGAAAC AAAAATGTCT GTCCGCATTG TCATCTTTGA	360
GAGCCTGAAC AAACCCCGAA CACCTCTGAC CTGGCACAGG GACCCTGAGT CTGCCTGCCT	420
GGCACTCTGA ACTTGCTGGC TACTGGGCTG AAGTCTCAGG ATCGAGGTAC AAGGACAAAG	480
CCTGTTGGTA TTATGTAGAC AGGAGCCCCA GGGGGCTCAG AGCAGGGGTT TACCGTTACT	540
CTTGGTCCAG CCGGCACAGT GGTCTTCCCT CAGTTTTTGT TTTTGT TTTT TGT TTTT	600
GTTTTGTTTT TGT TTTT TTTT TTTGTTGTT TGT TTTT TTTT TGT TTTT TTTT	660
CCTGCAGTCT GGACTGCCCC CTTGGGCAAT ACATTTTGAT GGAAACCAAC TTCTGTGTTC	720
CCAGATCGGC AGCCTCTCCA TGCCGAGCAC CCCTGTTCTA GCTAAGCCGG ACATGCTACA	780
CCCCTGCCGT GGTGGTTCA TCTCCCCACT CCCC GCCGTG TCTTAATAAC ACGGTGGATG	840
GATGGACAGA CGGACAGATG AAGCTTTGAC AGCTAGCCAC TGCTCTGCTC CTTTCTCTT	900
GTCTTCTGAG TAAATTCTCA AACTCAAGCC TCCAGGGGG AGCCAGAAC CCCACCCACC	960
ACTAGTGAAG AAGCAGCTTT TTAGGCAGCT CAGAGAGTGT CCGGAGCAGG GTTCCTACAG	1020
GCTCAGGTTG GGGGCAGTTT CAGTGTGTGG CTGCCGCCCTC AGGGCCCTCC CTGGTCAGCA	1080
CCTAGGTAGA GAGGAAGCAG GAAGGACCAA GTCCAGCTCA AAGGAACCAG TGAGGCTGGT	1140
TGAGAGCAGA CGTGGGTAT GGGTCTGTGA AAAGAAAAGT AGGCCGTTGG GAGCCCACCG	1200
GAGGACAGTC TGGCTCAAGG CATCACACCT GGTTCAGCA GCCTCCGAGA CAGCAGCCCCG	1260
TGGGTGTGAA CTTGGCTCCA TCTTGGCTGA ATTCTGCTGC CCCAGAGAAG AATCTGGCAC	1320
CGGAGTGGAC CCAGATACGT GTCACAGCCT TGGAATCG ATG GCA GCG AGA CCC	1374
Met Ala Ala Arg Pro	
1 5	
TTT ATT GGG CAC CTA GAG ATT TCC GAG AGG CGG TTC CTA ACA TTG TCA	1422
Phe Ile Gly His Leu Glu Ile Ser Glu Arg Phe Leu Thr Leu Ser	
10 15 20	
TTC CTT TCC CAC AGG CCT CAG GTC CAG GCC CAA AGC CAA GTG ACT GCC	1470
Phe Leu Ser His Arg Pro Gln Val Gln Ala Gln Ser Gln Val Thr Ala	
25 30 35	
CTC ATC GGC TCA GCA CAT CGG CGG GTG CTG GAT GTG CCG CCT GCA GAG	1518
Leu Ile Gly Ser Ala His Arg Arg Val Leu Asp Val Pro Pro Ala Glu	
40 45 50	
GCT CTT CGG GGC CAG GCT GGG ACT GGG CCC CTG AGA AAA CCC AAA CTC	1566
Ala Leu Arg Gly Gln Ala Gly Thr Gly Pro Leu Arg Lys Pro Lys Leu	

55	60	65	
TGT GCC TTA CCC AGT CAG GAT CCT TGG GAG CCT TGG CCC TCG ATG CCA			1614
Cys Ala Leu Pro Ser Gln Asp Pro Trp Glu Pro Trp Pro Ser Met Pro			
70	75	80	85
GGA CCT GGC CTA GCA TCA CCC TGG GCC ACC TGC ACG TTC ATG GCC TCA			1662
Gly Pro Gly Leu Ala Ser Pro Trp Ala Thr Cys Thr Phe Met Ala Ser			
90	95		100
CCA TCT GCT TTG CGA ACC ACT ACC CAC CCC CTG TCG CAT GCA CCT TCC			1710
Pro Ser Ala Leu Arg Thr Thr Thr His Pro Leu Ser His Ala Pro Ser			
105	110		115
ACC CCT GCG GTG TCT GAC TCT AGG GAG ATA GGT ACA GAT GTC TGT GAT			1758
Thr Pro Ala Val Ser Asp Ser Arg Glu Ile Gly Thr Asp Val Cys Asp			
120	125		130
GGG CTG AGA CAT CCC CCA CCC CCC ACG GCT CAT TGG TTT CTT TTA GCT			1806
Gly Leu Arg His Pro Pro Pro Pro Thr Ala His Trp Phe Leu Leu Ala			
135	140		145
GAG TGC TGT TTT CTG TGT CTT TAC TCT GCT ATG ATG ACT TGG GGG GTT			1854
Glu Cys Cys Phe Leu Cys Leu Tyr Ser Ala Met Met Thr Trp Gly Val			
150	155	160	165
AGT GAG GGC CTG AGG GAT GCT TGT TGG AAG TAC TGT GGG TAGGGCCTGG			1903
Ser Glu Gly Leu Arg Asp Ala Cys Trp Lys Tyr Cys Gly			
170	175		
TAGGTCCTTA GGGAAACAAAC AAACAAAAGG CTCATCTGCT CCTGGGTGGT GACCTGGCCA			1963
CAGCACCTGT CATCTCCCTT GGTGTCTGCC CAGTGGGGAC GAGCTTCCCT GGCTCAACCC			2023
CTATCAGGTG TGCACTGTTG AGTCTGCGGG CCAGCCGGGA CTTTAACCAA TAAAGAGCAA			2083
CCTTGTCTGT CAAAAA			2099

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 178 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Met	Ala	Ala	Arg	Pro	Phe	Ile	Gly	His	Leu	Glu	Ile	Ser	Glu	Arg	Arg
1				5					10					15	
Phe	Leu	Thr	Leu	Ser	Phe	Leu	Ser	His	Arg	Pro	Gln	Val	Gln	Ala	Gln
	20							25					30		
Ser	Gln	Val	Thr	Ala	Leu	Ile	Gly	Ser	Ala	His	Arg	Arg	Val	Leu	Asp
	35						40					45			
Val	Pro	Pro	Ala	Glu	Ala	Leu	Arg	Gly	Gln	Ala	Gly	Thr	Gly	Pro	Leu
	50					55					60				
Arg	Lys	Pro	Lys	Leu	Cys	Ala	Leu	Pro	Ser	Gln	Asp	Pro	Trp	Glu	Pro



65		70		75		80									
Trp	Pro	Ser	Met	Pro	Gly	Pro	Gly	Leu	Ala	Ser	Pro	Trp	Ala	Thr	Cys
			85						90					95	
Thr	Phe	Met	Ala	Ser	Pro	Ser	Ala	Leu	Arg	Thr	Thr	Thr	His	Pro	Leu
		100						105					110		
Ser	His	Ala	Pro	Ser	Thr	Pro	Ala	Val	Ser	Asp	Ser	Arg	Glu	Ile	Gly
		115					120					125			
Thr	Asp	Val	Cys	Asp	Gly	Leu	Arg	His	Pro	Pro	Pro	Pro	Thr	Ala	His
	130					135					140				
Trp	Phe	Leu	Leu	Ala	Glu	Cys	Cys	Phe	Leu	Cys	Leu	Tyr	Ser	Ala	Met
145					150					155					160
Met	Thr	Trp	Gly	Val	Ser	Glu	Gly	Leu	Arg	Asp	Ala	Cys	Trp	Lys	Tyr
			165					170						175	

Cys Gly

## (2) INFORMATION FOR SEQ ID NO:63:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1210 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW070
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: complement (639..928)
  - (D) OTHER INFORMATION: /label= SAC\_24286

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GTGACGCAG ATTTTGATGG CGTCCTTTAT CATATTTCAA ATCCTAATGG AGACAAAACA	60
AAAGTGATGG TCAGTATTTT TTTGAAATTC TACAAGGAAC TTCAGGCACA CGGAGCTGAT	120
GAGTTACTAA AGAGAGTATA CGGAAGTTTC TTGGTGAACC CAGAATCAGG ATACAATGTG	180
TCTTTGCTAT ATGACCTGGA AAATCTACCT GCATCCAAGG ATTCTATTGT GCATCAGGCT	240
GGCATGTTGA AGCGAAATTG TTTTGCCTCT GTGTTTGAGA AATACTTCCA ATTCCAAGAA	300
GAGGGCAAAG AAGGAGAGAA CAGAGCAGTT ATCCATTATA GAGATGATGA GACCATGTAT	360
GTGGAATCTA AAAAAGACAG AGTCACAGTA GTCTTCAGCA CAGTTTTTAA GGATGACGAC	420

```

GATGTGGTCA TTGGAAAGGT GTTCATGCAG GAGTTCAAAG AAGGACGAAG AGCCAGCCAC      480
ACAGCCCCCC AGGTCTTTT TAGCCACAGG GAACCTCCTC TAGAACTGAA AGACACAGAT      540
GCTGCCGTGG GTGACAACAT TGGCTACATC ACCTTTGTGC TGTTCCTCG CCACACCAAT      600
GCCACTGCTC GAGACAACAC CATCAACCTG ATCCACACGT TCCGGGACTA TCTGCACTAC      660
CACATTAAGT GCTCTAAGGC CTATATTCAT ACACGAATGC GAGCAAAAAC ATCCGACTTC      720
CTTAAGGTGC TCAACCGTGC ACGCCCAGAT GCCGAGAAAA AAGAAATGAA AACAATCAGC      780
GGGAAGACTT TTTATCCCG CTAACCTTG GGAAGTGGG AGGAAGCACT GTTGACCGAA      840
GGCTGGACCG CTTGCTACTG GATAATCGTA GCCCTTCATG TTGCACCTTC TCCAGGTTCT      900
TAAGGGATTC TCCGTTTGGG TTCCATTTTG TACGAGTTTG GCAAATACGC TGCAGAAACG      960
AGCTGTGCTT GCAAGGACTT GATAGTTCCT AAGAATTAAA AACTCACTTG ATCAACTTAA     1020
TTCCCTTTTA TTTTCTCTCC CTCACCTCCC TTCCTTTTCC AAGCTGTTTG CTTTGCAATA     1080
TGTTACTGGT AATGAGTTGC AGGTAATGCA ATCTTAACCT GTTTTCTTCT AAGTATTTGA     1140
GTTCAAAACT CCTGTATCTA AAGAAATACG GTTGGGGTCA TTAATAAAGA AAATCTTTCT     1200
ATCTTAAAAA                                     1210

```

## (2) INFORMATION FOR SEQ ID NO:64:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1844 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW073

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 9..1223

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 598..1036  
 (D) OTHER INFORMATION: /label= SAC\_24456

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

```

GTCGACCT GCC GGC CGA CAG GGC CGC GGG ATG AGG AAG CGG ACC GAG CCC      50
  Ala Gly Arg Gln Gly Arg Gly Met Arg Lys Arg Thr Glu Pro
    1             5             10

GTC ACC TTG GAG CAT GAG CGC TGC GCC GCC TCA GGC TCG TCT TCC TCC      98
Val Thr Leu Glu His Glu Arg Cys Ala Ala Ser Gly Ser Ser Ser Ser
  15             20             25             30

```

GGC TCG GCC GCC GCG GCG CTG GAC GCC GAC TGC AGC TTG AAG CAG AAC Gly Ser Ala Ala Ala Leu Asp Ala Asp Cys Ser Leu Lys Gln Asn 35 40 45	146
CTG CGT CTG GCG GGC AAG GGG ACG GCA GAG CCG CAC AGC GCA TCC GAC Leu Arg Leu Ala Gly Lys Gly Thr Ala Glu Pro His Ser Ala Ser Asp 50 55 60	194
GCG GGC ATG AAG CGG GCT CTG GGC AGA CGG AAG AGC CTG TGG TTC CGA Ala Gly Met Lys Arg Ala Leu Gly Arg Arg Lys Ser Leu Trp Phe Arg 65 70 75	242
CTA AGG AAG ATA CTT CTC TGT GTT TTG GGG TTC TAC ATT GCC ATT CCA Leu Arg Lys Ile Leu Leu Cys Val Leu Gly Phe Tyr Ile Ala Ile Pro 80 85 90	290
TTT CTT GTC AAA CTG TGT CCT GGG ATA CAG GCC AAA CTG ATA TTC TTA Phe Leu Val Lys Leu Cys Pro Gly Ile Gln Ala Lys Leu Ile Phe Leu 95 100 105 110	338
AAT TTC GTG AGG GTT CCC TAT TTC ATT GAC TTA AAA AAG CCA CAG GAT Asn Phe Val Arg Val Pro Tyr Phe Ile Asp Leu Lys Lys Pro Gln Asp 115 120 125	386
CAA GGT TTG AAT CAC ACC TGC AAT TAC TAC CTC CAG CCC GAG GAT GAT Gln Gly Leu Asn His Thr Cys Asn Tyr Tyr Leu Gln Pro Glu Asp Asp 130 135 140	434
GTC ACT ATT GGA GTC TGG CAC ACC ATT CCC TCT GTC TGG TGG AAG AAT Val Thr Ile Gly Val Trp His Thr Ile Pro Ser Val Trp Trp Lys Asn 145 150 155	482
GCC CAA GGG AAG GAC CAG ATG TGG TAT GAG GAT GCT CTG GCT TCT AAC Ala Gln Gly Lys Asp Gln Met Trp Tyr Glu Asp Ala Leu Ala Ser Asn 160 165 170	530
CAC CCC ATC ATC CTG TAC CTG CAT GGG AAT GCA GGC ACC AGA GGA GGT His Pro Ile Ile Leu Tyr Leu His Gly Asn Ala Gly Thr Arg Gly Gly 175 180 185 190	578
GAC CAC CGT GTG GAG CTG TAC AAG GTG CTG AGT TCC CTT GGT TAC CAC Asp His Arg Val Glu Leu Tyr Lys Val Leu Ser Ser Leu Gly Tyr His 195 200 205	626
GTG GTC ACC TTC GAC TAC AGA GGT TGG GGT GAC TCA GTA GGA ACA CCA Val Val Thr Phe Asp Tyr Arg Gly Trp Gly Asp Ser Val Gly Thr Pro 210 215 220	674
TCA GAG CGA GGC ATG ACA TAT GAT GCA CTC CAT GTT TTT GAC TGG ATC Ser Glu Arg Gly Met Thr Tyr Asp Ala Leu His Val Phe Asp Trp Ile 225 230 235	722
AAA GCA AGA AGT GGT GAT AAT CCT GTG TAT ATC TGG GGC CAT TCG CTG Lys Ala Arg Ser Gly Asp Asn Pro Val Tyr Ile Trp Gly His Ser Leu 240 245 250	770
GGC ACT GGA GTG GCA ACA AAT CTG GTC CGG CGC CTT TGT GAG CGA GAG Gly Thr Gly Val Ala Thr Asn Leu Val Arg Arg Leu Cys Glu Arg Glu 255 260 265 270	818
ACG CCA CCA GAT GCC CTT ATA TTG GAG TCT CCG TTC ACA AAT ATT CGT Thr Pro Pro Asp Ala Leu Ile Leu Glu Ser Pro Phe Thr Asn Ile Arg	866

275	280	285	
GAA GAA GCA AAG AGT CAT CCA TTT TCA GTG ATA TAC AGA TAC TTC CCA Glu Glu Ala Lys Ser His Pro Phe Ser Val Ile Tyr Arg Tyr Phe Pro 290 295 300			914
GGC TTT GAC TGG TTC TTC CTC GAC CCC ATT ACA AGC AGT GGA ATT AAA Gly Phe Asp Trp Phe Phe Leu Asp Pro Ile Thr Ser Ser Gly Ile Lys 305 310 315			962
TTT GCA AAT GAC GAA AAC ATG AAG CAC ATC TCC TGT CCC CTG CTC ATC Phe Ala Asn Asp Glu Asn Met Lys His Ile Ser Cys Pro Leu Leu Ile 320 325 330			1010
TTG CAC GCT GAG GAT GAC CCA GTT GTA CCC TTT CAT CTC GGC AGG AAG Leu His Ala Glu Asp Asp Pro Val Val Pro Phe His Leu Gly Arg Lys 335 340 345 350			1058
CTA TAC AAT ATT GCT GCG CCA TCC CGA AGT TTC CGA GAC TTC AAA GTC Leu Tyr Asn Ile Ala Ala Pro Ser Arg Ser Phe Arg Asp Phe Lys Val 355 360 365			1106
CAG TTC ATC CCC TTT CAC TCA GAC CTT GGC TAC AGA CAT AAG TAC ATC Gln Phe Ile Pro Phe His Ser Asp Leu Gly Tyr Arg His Lys Tyr Ile 370 375 380			1154
TAC AAG AGC CCA GAG CTT CCA CGG ATA CTG AGG GAA TTC CTA GGG AAG Tyr Lys Ser Pro Glu Leu Pro Arg Ile Leu Arg Glu Phe Leu Gly Lys 385 390 395			1202
TCG GAA CCA GAG CGC CAG CAC TGAGCCTGGC CCGTGAAGGA GCATGGAGAC Ser Glu Pro Glu Arg Gln His 400 405			1253
CCACCTTCCT TCCCTTCTCC CTGAACAGCA GTCTGGCACC CAGAAGCTCA GAGTGCCACC			1313
ACCTGTGGTG CTCAGGAGCC CAGCCTAGAA AGAGGACTCC GACACAGCGG GCAGAGGCTC			1373
CACAGACGGA TCTATGAGGA AAATACGGTG GCAGGCAGGC AGGCAGGCGA CCCCCTGACC			1433
CTCTGGTGGC CGCTGTATCT GAGCCCTTTT GGGAAGGCTT ATAGACAACA GGTGGAGCCC			1493
ATACGCTGGG CATAGGGAGC CTGGGAAGGG CTCAGGAGCT CAGGACCACT CCAGGCTCTC			1553
TAGCACCACC GCTTAAAATA CAGGAAAAAG GTTCTTTCTG CCCTTCCTGG CGTACACAGA			1613
ACAGATTCCA AGTGGTTCAA TTTGTCCCCT ACAGCTCATG TACCTGCTTG CCTTCCTCAG			1673
CTGTCCCTGC CTCTCCTGGC ATCTGTACAC CCACAGTGAG GGGCACCTGG ACTTGCACCT			1733
CCATTCTGCC CACCTGTCTG TCACCTAACC TGGCCGTAGA CTGAGCATTT ATTTAAGAAT			1793
AAAATCTCGG TGGTGGTCTA TTTGTTTTTT TCTACAAAAA AAAAAGTCGA C			1844

## (2) INFORMATION FOR SEQ ID NO:65:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 405 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Ala Gly Arg Gln Gly Arg Gly Met Arg Lys Arg Thr Glu Pro Val Thr  
 1 5 10 15  
 Leu Glu His Glu Arg Cys Ala Ala Ser Gly Ser Ser Ser Ser Gly Ser  
 20 25 30  
 Ala Ala Ala Ala Leu Asp Ala Asp Cys Ser Leu Lys Gln Asn Leu Arg  
 35 40 45  
 Leu Ala Gly Lys Gly Thr Ala Glu Pro His Ser Ala Ser Asp Ala Gly  
 50 55 60  
 Met Lys Arg Ala Leu Gly Arg Arg Lys Ser Leu Trp Phe Arg Leu Arg  
 65 70 75 80  
 Lys Ile Leu Leu Cys Val Leu Gly Phe Tyr Ile Ala Ile Pro Phe Leu  
 85 90 95  
 Val Lys Leu Cys Pro Gly Ile Gln Ala Lys Leu Ile Phe Leu Asn Phe  
 100 105 110  
 Val Arg Val Pro Tyr Phe Ile Asp Leu Lys Lys Pro Gln Asp Gln Gly  
 115 120 125  
 Leu Asn His Thr Cys Asn Tyr Tyr Leu Gln Pro Glu Asp Asp Val Thr  
 130 135 140  
 Ile Gly Val Trp His Thr Ile Pro Ser Val Trp Trp Lys Asn Ala Gln  
 145 150 155 160  
 Gly Lys Asp Gln Met Trp Tyr Glu Asp Ala Leu Ala Ser Asn His Pro  
 165 170 175  
 Ile Ile Leu Tyr Leu His Gly Asn Ala Gly Thr Arg Gly Gly Asp His  
 180 185 190  
 Arg Val Glu Leu Tyr Lys Val Leu Ser Ser Leu Gly Tyr His Val Val  
 195 200 205  
 Thr Phe Asp Tyr Arg Gly Trp Gly Asp Ser Val Gly Thr Pro Ser Glu  
 210 215 220  
 Arg Gly Met Thr Tyr Asp Ala Leu His Val Phe Asp Trp Ile Lys Ala  
 225 230 235 240  
 Arg Ser Gly Asp Asn Pro Val Tyr Ile Trp Gly His Ser Leu Gly Thr  
 245 250 255  
 Gly Val Ala Thr Asn Leu Val Arg Arg Leu Cys Glu Arg Glu Thr Pro  
 260 265 270  
 Pro Asp Ala Leu Ile Leu Glu Ser Pro Phe Thr Asn Ile Arg Glu Glu  
 275 280 285  
 Ala Lys Ser His Pro Phe Ser Val Ile Tyr Arg Tyr Phe Pro Gly Phe  
 290 295 300  
 Asp Trp Phe Phe Leu Asp Pro Ile Thr Ser Ser Gly Ile Lys Phe Ala  
 305 310 315 320

Asn Asp Glu Asn Met Lys His Ile Ser Cys Pro Leu Leu Ile Leu His  
 325 330 335  
 Ala Glu Asp Asp Pro Val Val Pro Phe His Leu Gly Arg Lys Leu Tyr  
 340 345 350  
 Asn Ile Ala Ala Pro Ser Arg Ser Phe Arg Asp Phe Lys Val Gln Phe  
 355 360 365  
 Ile Pro Phe His Ser Asp Leu Gly Tyr Arg His Lys Tyr Ile Tyr Lys  
 370 375 380  
 Ser Pro Glu Leu Pro Arg Ile Leu Arg Glu Phe Leu Gly Lys Ser Glu  
 385 390 395 400  
 Pro Glu Arg Gln His  
 405

## (2) INFORMATION FOR SEQ ID NO:66:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 855 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

(B) CLONE: HW074

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 7..300

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION: complement (64..365)  
 (D) OTHER INFORMATION: /label= SAC\_24464

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GTCGAC AGC CCT CGA GGG AGC AGA CGC CGC GGG CCA CTG GCT TCC TCT	48
Ser Pro Arg Gly Ser Arg Arg Arg Gly Pro Leu Ala Ser Ser	
1 5 10	
ACT CAG GAA CCA CGG ATG TCA CCG CCC AGT CAG CTG TGT CTC CTC ACC	96
Thr Gln Glu Pro Arg Met Ser Pro Pro Ser Gln Leu Cys Leu Leu Thr	
15 20 25 30	
ATT GTC GCC CTG ATT CTG CCT AGT GAA GGG CAG ACA CCA GAA AAA CCC	144
Ile Val Ala Leu Ile Leu Pro Ser Glu Gly Gln Thr Pro Glu Lys Pro	
35 40 45	
AGA TCC AGT TTT ACA GCG CAC CAG AGT TCT GTG ACT ACT CAT GTC CCA	192
Arg Ser Ser Phe Thr Ala His Gln Ser Ser Val Thr Thr His Val Pro	
50 55 60	
GTT CCA GAT CAA ACT AGC CCA GGA GTC CAG ACC ACT CCT CCC ATC TGG	240

Val Pro Asp Gln Thr Ser Pro Gly Val Gln Thr Thr Pro Pro Ile Trp  
65 70 75

ACC AGT GAA GCT GCG AAG CCA CAG GAA GCC AGA CAG CAG CCA AAA CCA 288  
Thr Ser Glu Ala Ala Lys Pro Gln Glu Ala Arg Gln Gln Pro Lys Pro  
80 85 90

AGA CCC AGC AAC TGACCGAAAT GGCCACTGCG AATCCAGTGA CAGATCCAGG 340  
Arg Pro Ser Asn  
95

GCCACTTACA AGCAGCGAGA AAGGTACCCC GGCACCTCTCC AGGATCAAAT CTCCCAGCCC 400

ACCCAAAGGT TACATGCCTC CATCCTACAT TGAGAATCCA CTGGATCCCA ATGAGAACAG 460

CCCCTTCTAC TACGACAATA CCACCCTCCG GAAACGGGGG CTGCTGGTGG CGGCAGTGCT 520

GTTCACTACT GGAATTATCA TCCTCACTAG TGGGAAGTGT AGACAGTTCT CTCAGTTATG 580

CCTGAATCGC CACAGGTGAG TGGGAGCCAG CACCCTGATG GGCACCCCAA CTGGAGCCGC 640

CATACCATAC CAGTTCACCA CCCCTGCCTC CCTCCCTCTG CTCCAAGAGC CAACAGAGTG 700

GTCAACATAA ATGGATCCTC AAAGGAAGAG GCCACCGGAG GGAGCCAGGC CTAAGGCTAA 760

ATGGTCTTCC CACCCTGAGG AGAGAGGTCT CCCAGGCAC TGCTGTGATC CTGCCTATCC 820

TGTTAGATA AATCCACATG GTCTCTCTTC AAAAA 855

## (2) INFORMATION FOR SEQ ID NO:67:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 98 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Ser Pro Arg Gly Ser Arg Arg Arg Gly Pro Leu Ala Ser Ser Thr Gln  
1 5 10 15

Glu Pro Arg Met Ser Pro Pro Ser Gln Leu Cys Leu Leu Thr Ile Val  
20 25 30

Ala Leu Ile Leu Pro Ser Glu Gly Gln Thr Pro Glu Lys Pro Arg Ser  
35 40 45

Ser Phe Thr Ala His Gln Ser Ser Val Thr Thr His Val Pro Val Pro  
50 55 60

Asp Gln Thr Ser Pro Gly Val Gln Thr Thr Pro Pro Ile Trp Thr Ser  
65 70 75 80

Glu Ala Ala Lys Pro Gln Glu Ala Arg Gln Gln Pro Lys Pro Arg Pro  
85 90 95

Ser Asn

## (2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4315 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW075
- (ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: complement (3785..4014)  
 (D) OTHER INFORMATION: /label= SAC\_24466

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

GTGACATGG CGAAGTACTC TAGCCTGGAA GCCAAGCTCT GCCAGGTTGA AAGTAAATAC	60
TTGATACTGC TCCAAGAAAT GAAGACACCA GTTTGCTCAG AAGAACAGGG CCCTGCCAGG	120
GATGTCATAG CCCAGTTGTT GGAGGACGCT CTGCAGGTTG AGAGCCAAGA GCAGCCAGAG	180
CAAGCGTTTG TCAAGCCTCA TCTGGTCAGT GAGTTCGATA TCTACGGGTT TAGGACTGTC	240
CCTGACGATG ATGAGGAAGA GAAATTGGTC GCCAAGGTCC GAGCATTGGA CCTGAAGACT	300
CTGTACCTCA CAGAAAACCA GGAAGTTTCC ACTGGGGTCA AGTGGGAAAA CTATTTTGCA	360
AGCACAATGA ACAGGGAGAT GGTGTGCTCT CCTGAGCTGA AAAACCTGAT CCGAGCAGGC	420
ATTCCCATG AGCACCGCTC CAAGGTGTGG AAGTGGTGTG TTGACCGTCA CACCAGGAAG	480
TTCAAGGACA GCATGGAGCC AGGCTACTTC CAGGCCTTAC TCCAGAAGGC TCTAGAGAAA	540
CAGAACCCGG CCTCCAAGCA GATTGAGCTG GACCTGCTTC GGACTCTGCC CAATAACAAA	600
CATTACTCCA GCCCCACGTC AGAGGGCATA CAGAAGCTTC GCAATGTCCT GCTTGCCCTTC	660
TCATGGAGGA ATCCGGATAT TGGCTACTGC CAAGGCCTAA ACAGGTTGGT GGCAGTGGCG	720
CTCCTTTACC TGGACAAGA GGATGCTTTC TGGTGTCTCG TTACCATTGT GGAAGTCTTC	780
ATGCCTCGAG ACTATTACAC AAAGACTCTA TTAGGATCCC AGGTGGACCA GCGGGTGTTT	840
AGAGACCTCC TGAGTGAGAA GCTGCCTCGA CTGCACACCC ATTTTGAGCA GTACAAAGTG	900
GACTACACCC TCATCACCTT CAACTGGTTT CTGGTGGTAT TCGTGGACAG CGTTGTCAGC	960
GACATCCTCT TTAAGATATG GGAATCTTTC CTTTATGAGG GACCAAAGGT TATTTTCCGT	1020
TTTGCCCTGG CACTTTTAA ATACAAGGAA GAGGAGATCC TGAAGCTGCA GGATTCGATG	1080
TCCATTTTCA AGTATCTCCG ATACTTCACT CGCACTATCC TTGATGCCAG GAAGCTGACC	1140
AGCATCTCCT TTGGGGATCT GAACCCCTTC CCCCTGCGCC AGATCCGGAA CCGGAGAGCC	1200



TACCACTTGG AGAAGGTCCG GCTGGAGCTG ACAGAGCTGG AGGCCATTCTG AGAGGACTTC 1260  
CTGCGTGAGC GGGACACTAG CCCTGACAAA GCGAGCTGG TTAGCGATGA GGAGGAAGAC 1320  
ACTTGAATGA ATCCCCACCC CAGGATGCTG CTTTCCTTCT TGCCTTCCTT CATAACCAAA 1380  
AGTAGCAAAG GCGCATCTGC ACATCAGCTG TCAGAGTGAA ACATCTACAG CCTCTGGGCA 1440  
CCTGCGGGGC AGATGTCCAA AGCTAAGCTG CACTTTCTCG TCTGCCATGA GGGCAGGGGT 1500  
CTGATGGCTT TACAGCCATT TCCATGCCGT GGCATTCACT CATATGCGTC ATATGCTTGC 1560  
TGCCTGATGG GTGAGTCACC AGGCGACTGT TTCCAAGGGC CTGTTTACAT CATCTGCCAG 1620  
TTGTATGTTT CACCTACAGA AACTGTTTGA CATTTCCCTG TGGCTTTCT TTTTATCCAA 1680  
GGGCAACTCT GTGGAGGAGT GCACATGGAG CTGCAGCGTG GCTGCATCCA GGGGTAAAAG 1740  
GTCACCATAA GAAGAGGGAC CTCTTTAAAA AAAAGCGAGC ACAGAACTGG CCTCCCTAGA 1800  
GAGCGTGGGA GCCCTCTGCT GGTGGCTTCG CACTGCCAGC CACTATCCCT TCTGGACTAA 1860  
CCTTTGCCCT CCTATACAGC CTTTTCAGAT TGTAGCACTC ACGCTGCTCT GCAGCCAGCC 1920  
CCACACTTCT CATTGCCAGC TATTCTGCCA GATGTCTGTC CCAGCTTACA GTGGGGCTAG 1980  
CGCATCTGGG AGGTGACTTA TAGACAACAT TCTCCAACCTG TGGCTAGATT GGCCTTCCAA 2040  
AGGAACCTTA CAGTAACCCA CCCCTGCATA AAGGAACACA GGAAACTCCT ATCTGTGGCT 2100  
TGTGGGCTCC TGGATGGCAG TGGCCAGGGC TCTAGAATCT CTTCTGCCCA ACAGCTGGGT 2160  
GGGGAGTAGA GCTGACTGCC AGCTTCTTAC ATCTGGAAGC ACCTTTGCTT GGAACAAGA 2220  
GTAATGAGTT CTCCTAGTCT CTGTGTTGCC TGGATCATCC TTGTCCTTAA CAGGGATCTG 2280  
AGGACACTGT GGCTCTGGT ATAGCAGACT CTGGGGACTT GGAACACTTC CTTCTTGAAG 2340  
CCCAGAGAGG GTCTGAACAA AAAAGGTTAT TTTTGCTCAG TCCCAGAGCA TCCCTAATGG 2400  
TGGCTGTTAT GAGAAAACAT CTTCTTACAC AGTATTATGG GAACACAAGG GAAACAGCCT 2460  
CAAGTGGACT ACTTTACACT ACCTAGGGAA AAGAAATTAG AAACAATCCT TAGTGTAAC 2520  
ACACTCAGAC TGAGGAGTGG AGCAAGACAG TTTTCGAGAA CGCTTTTTC ATGACCCGTT 2580  
TTCTGTGTTT TGCTGTCTGC TTTCTAGCAC TTGAATGGCT GAGTCCATTT TCAGCATCTA 2640  
AATTTGCTTC ACTTCTCTAA AGTTGCTTCA TATATTACCA TGTAAGGA GCTGGTCTCT 2700  
ATATTTAGCC AATAGCCCGT CTCTTTTCAG GTATCAGAAA TGTCATCCTT GTGTCTTTGA 2760  
GATGATGGTT TGAGGCCCCG CTTCTTGAAG AAGCTGTCTG GAGCTGGACA GTGATCTATC 2820  
TCACCACTGT TCCTAGGCCT CCCTGAGACC AAATGCTGGA AGATGCCCCA CCCCCTGTAA 2880  
GATGCCGGCT GTCAGCCTTG GGGACAGCTG TTAGGAGGTT AGAGTTGGAG CCCCATGGCT 2940  
TAGCTCCAGC CGCTCTGATT TCGGTTGTCA TGTCAATTTT CTACCAGAAG ATCCAGTCCA 3000  
AATAAGTATT TGGTGTGGAT GCTCTCAGTG GAGCATTTGG TGTGGATGCT CTCAGTGGTG 3060  
GGTTGCTTTC GCCCTGGCAC CAGTGGAGGC ATCTTTAAGT ACTTGGTGAT GCTGAGGACA 3120

```

GGACCATGTT CCTGGCATGC GGGGCCCTGG GTTCCATCCC AGTGATGAAA AAAAATCTTC 3180
AGTATGTGGG GATTGTGGGG AGGCGCTGTT CTGCCATCCC CTGAGAGACA GCATTAGTTT 3240
CTCAGGCAGT TAGTATCAGT GCCCAGGTAA AAGTCTCAAG AGGTCCCAAA GGGACTCTCC 3300
TCCCTGGAAT AAATGACTCG ATGTCTGCAG AGCGGAGAAG GGACGGTTGC CCTCTTTCCA 3360
GTTTGGGGAC AGAATAACAG AGATTCATCC CAGAGGCACA GCCCCAGTGG TTGAGCCATC 3420
GTCTCTCATG CCCTCGTCAC TGAGATTGTC TTTACCTCC AGTACTCCTG AGACTCCAGA 3480
CTTCTACCAG ACACCATGGC AGGTACCCCT AAGTCACTG GCTGCCTCAG GCTGCTTTTC 3540
ATCTAAACCT TTACCCAGCA TGCAACCCAT GGCTCCTGCC TGGTGAGCCT TGTGGGCTGA 3600
GAAACCCAAA GCACAGCCTC CAGCGCGTGG GTCCAAATGC GTGTGCCACA GTAGGGGAGC 3660
ACCTTGCCCT CAGCACCCCA TATGGTGGGA CTTGCCAGTA TTTCAACTGT TGATACTTTT 3720
ATATTTTCAC TACATTTTAT ATTTTATAG ACTATTTTAT ATTAAGGTTT TTTCTAGATT 3780
CTGTACAACT TTTATATAAC AAGTTCATT TGTGTGCAC ATCTCCCCTG TATGTATTTA 3840
TAGCAACCCA AGCCTCTCAA ATTCCAGGTC CCTCTGTGC CACCTTTTCC CCTTAGCCTT 3900
GACTGTCACT CTCTTCTGCT CACCGAGGTA GCCTCGGGGG CTACTTCTAG CCACCATGAC 3960
AGCCAGACAG CAGGAACCTG TGTGTGTCAG CTGTGGTCTG TTGATACAGG AAGTACTCTG 4020
GGCAGTAACC CAGTAGCTAT GTTACTGCAC AAGCCCTTCT CGCTTAGTGC AGTGATGTCC 4080
TGGTGAGGCC ACCTGTGGGC GGGTGGCTGG ACCCATCACG AAAGTCCTTG AAGACAGAGG 4140
CTAGATGAAA GCCCAGTGTC TGAGGTCCTT CTGTGGGCAA CTGGAGGACT TGAGGGCACA 4200
CTGCCAAGTG ACTGTCCTTT TATATCCTAG CAGCTTCTGT GCTCACCTT GAGTTAAATG 4260
TTAGACTTTT AAAGGAAGAA ATAACTAAT TTGTATTACA TTCTGCCTCG AAAAA 4315

```

## (2) INFORMATION FOR SEQ ID NO:69:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1668 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW076

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (42..281)
- (D) OTHER INFORMATION: /label= SAC\_24409

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

GTCGACTTTT GCCTTTAAAA TTCTGTTTAC TTAGCATATG TACACGTAAA AGAGAAATCA	60
CCCTTAGATC CCACCCAGCA GAAACATGC ACAAACCCAA GGTATATGTG AAAGGGAGGG	120
CCCCAGCCCC AAGTTAGCAG TTGGGATGGT GCAATCCTGT GAGTCACAAG CCCATGGCAA	180
GAGGGCTGGA CTAAAGTTTC AACACAGCCA GAACCCCTGC TGTGGCCCT AAAATAAGCG	240
TGCACCATAG CCAAAGGTCT GTTTGATGCC CTCAAAGTAG TACCGCTGCC AGCCTTGCCG	300
TGTCCGCTCT TCCTCGGGAG CAGGGATGCC GCGGCCCTCC ATGCACAGCT CTGTCTCTCC	360
ATTCTTATCA ATGAAGGTCA AGGTGATGGT CGCGAAGTGC CCCTCTGGCC ATGACTTAAA	420
CCTCCACTTC ATAGCAATAT GTTTCTCAGG AACCAAGTCA GTAAACTCCC CAGTGACGTT	480
ACCATCGACC ATGTGAAACT TCCCTCCTCT GTCTGCTTCC AAGGCCGCAG GAGCATGGGT	540
AAAGGCCTGG ACCAGCTCCT GCGTGGTGAA CACTCTGTAG AGCTCCTCTG GAGAGGTCAG	600
GAAGGTTTCT TTGAGGGTGA TCTTACAAGT GGGTATTTTG ACACCAACAG GTTTGGCCTG	660
GCTTTTGTAA GGAGCACACT TAGCCTTGCA CACTTCCGTT TTTAGTGCTG GTTGGCCAC	720
TGGGTCTACT GATTCTCCAT TCACTGTGGG CAAGATCATG CCCTGAGTAA ACTCTGTCTT	780
GAGGGTGCTG ATGTAAATTC CCACTGCTTC TCTGAGAAGT TTCACCCCGT CTTCTTTCAT	840
TAAGGCCACG AACTTGTGT CAGGCTCATC TTTGGCAAGG CTGCTTCTTA GAGCCCTTGT	900
TATGGCAGCG CTCCATCTAG GAAACAGCAG TCAGTCTCTC CTGCTCACTG TTCTCCCTCT	960
CTGTGCTCAA GTCTGTCAGT GCTCGCTGGC CTCTGGCTCT TCCCTTTGGG GAGGGAGCGG	1020
CTGTTCTGCC TGCTTCACCG AGTCTTCACA AGCTGAGGGG ATGAGGCCAG GTTAGTACCA	1080
GGGACCAGCA ACCTTTCAGG GTCCTTGGG GGTCTCAGGT CCACACACAG TGAGCATCCC	1140
TTGCCAGGTG CCCCATGTAT CTGTCAGTGC TGAGACGCCA AACACACAAC CTGTTCCAG	1200
CTGCTGCTGG GGGCCACTGT GACCGAGGAG AGTGTGGGGT TGCCAGGGAG AGCGCTGTCC	1260
CTCTGACAGC CTGACCAGAT GATGTAGCTT TGAAAGCCGG CAGCATGTGA GTAGCAAGAT	1320
GTGACCTTGA GCTAGCTGAG CAGAACACAG GTGGATGGGC AGAGGCCCCC ATCCTGGCGA	1380
GCTGAGCCTG GCTGTGTCCT GTCAGGACGC CCTACTCTGT GACCCTACCC TGTGATCTGT	1440
GAAGTGGTGC AGTCGGAAGA CGAGGCTGGA GTCCTTTACA AGTTTCTACA CTAGATTGTA	1500
GCCAGGTTTT CTCCAGGCC TTAGGCTTTG GTTACTTTT AATTTTAATT TCACTGTTAC	1560
CCTTGACTAT TGTCTTTTT GTTGCAGTGT TGGAGAGGCA GGAGGTATTT GTGCCTCATC	1620
TATTACTGCA AAAAAATTAA ATAAATGTAA CAATAACGT CCTAAAAA	1668

## (2) INFORMATION FOR SEQ ID NO:70:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1019 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW078

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (211..743)  
(D) OTHER INFORMATION: /label= SAC\_24345

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GTCGACGCCA GCGTGAGCGC CCTGCAGCGC CTGGTGGAGC AGCTGAAGCT GGAGCCCGGC	60
GTGGAGAGGA TCAAGGTCTC TCAGGCAGCT GCAGAGCTTC AGCAGTACTG CATAAGAAT	120
GCATGCAAGG ACGCCCTGCT GCTTGGTGTT CCGGCTGGAA GCAACCCCTT CCGGGAGCCC	180
AGGTCTCTGTG CTTTACTTTG AAGACTGGGA AGAAGCTCTC TGAGGAACCT GTGTGTGCAT	240
GGTGATGGAT GACTACCGCC AAGTCCCAAG AAAACAGATT TCAAGAGCCA AATCATTTTC	300
TCTGTGGCCT GGTCTTAGGT CAAAATCTA CAGAAATCAA TGTTTGTCT TATTTTTTAA	360
CCTCATTACC AAACCTACTG GGTGCATCTC TCAGTCAAGC CTGTGCGGCC CTGCCTGGGA	420
TGGAGGCATG TTTTGTTTAC TGGGTATTTG CATTTCTGGA TGGAAATGTT GTATTTCCGA	480
AAACTGAGTA TTTGTAATTT TCCTAACATT TTTTACTGT AATAAAAAAT TTTCATAAAA	540
ATTTACAGTT ACTAAATTAG ACAAAGTATA AATTGGGGAA TGATTGATAT GCTGGTTTCG	600
TAAAGGAGGC CTTTCCAGTC CCAGTGAGCA AACGGTCTGA CCTGGACGGG AAGCTGGGGA	660
GCCGCTCCTT AGCACACGGA GCATGGATGT GTTGTTAGGT GGCCCTCCCG TGTGAAGCCA	720
TTAGAACTTT CTTTAGTTCC TGTA CTCTAC AACTTATAAT TGTGTCTGAT GCTTTGAAAA	780
GTGTGCTCAT GCCTTTAAAA ATGTTTAAAT GTGATCTACA CTTAGAACGT TGCAATTGGT	840
TAAAAACAAA ATATGAAGAT ATTATAACCT AAAAGAATTC ATGTCACACA TATTTTATTG	900
GGGGTGATGT GCCTTTGATT TAATTGGGGA CACTTTTAGA AGGAAAGTTT GTGTTTGTGA	960
TTATCTTTGA AGAGCTTGGA AATAAAATTT TTGCCTAATT CATCATTTTC CATGAAAAA	1019

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 961 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: HW080
- (ix) FEATURE:
  - (A) NAME/KEY: misc\_feature
  - (B) LOCATION: 629..919
  - (D) OTHER INFORMATION: /label= SAC\_24033

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

```

GTCGACGCTG GGC AAAAGAA GGTTCACCA GAGCCTGTGG GTTGTCTCCTT AGGCTAGAGG      60
CCCCAAGATT TGGAACTTGG TGAATCAGT AGGAGGGACA TTGAAGCTCA CAGATATATC      120
ACACGTTGGT CAGTTGCTTC TGCAACTGGT CCGTTAGGTG CTGTGGGACG AGCCCTTTTC      180
TACCTTCTTC AGCTTCTACA TCCACTGTCT GCCTGTGTCT CAGTTTACAT CTTCTTTTG      240
AATATGGAAT CTCAATAGC CCAGGCTAGC TTCAAATTG TTAAGTAGTT GAGGACAACC      300
TCAATCTTTC TATTCGCTGC CTCCACCTCT CTCAGTTTAC CTTTTTTTTT TTTCTCCTTA      360
GGATTAAGTC CCGTACAGGC CCAGAGTGGT AAGCCATAAC ACCCCTGGTC TTTCTCTCTT      420
CCTCTCAAGA TTTCTCAGG CTACCCCTTT TCCTTCTAGC TCTCTCTCTC TGCTAACGCC      480
GAGCCCTGAT TGTTAACTG TGTCTCCCTC TTCATCCTTC TGAGACAATT ACCCAGGATG      540
CGAATGTTCC TCTGTGAGCC CGGTGTACT GGCTGGGATT GTGCTGGGCG ACTTGGTGCT      600
GACTCTGCTC ATCGCCCTGG CGGTGTACTC TCTGGGCCGC CTGGTCTCTC GAGGCCGAGG      660
GACTGCAGAC GGGACCCGGA AACAGCACAT GGCTGAGACT GAGTCACCTT ATCAGGAGCT      720
TCAGGGTCAG AGGCCAGAAG TATACAGTGA CCTCAACACA CAGAGGCAGT ATTACAGATG      780
AACCCACCCT ATGCCCACCA ACAACCTGAT GCGCGGATCC ACTCATTCCTA GACGCTTACT      840
CAACAAACCC TCCCTGGGAT CAGGACTCCC GCTGGAATAA ATATCCACAG AGTGCCCTCT      900
GGGAGATATC TGACCTTGTA CCGTTTCTGT CCCCATAA AAGACGAAGC AAAAACAAAA      960
A                                                                                   961

```

## (2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 1890 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW082

(ix) FEATURE:  
(A) NAME/KEY: CDS  
(B) LOCATION: 150..569

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: 1381..1848  
(D) OTHER INFORMATION: /label= SAC\_24469

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

GTCGACCGCG GCCGCGTCGA CGGACTTGCG GTTCCTTCAG CATTCTGCCC GCGCACTGAG	60
TGGCTGTTTC CTGGAGCTAC TGTGTTCCCT CGGGAGTGCA GCAGATCCCG TTCAGAACCT	120
ATTACCAAGA GCCCCGAGCA GGACCCACC ATG CAG TGC TTC AAA TTC ATT AAG	173
Met Gln Cys Phe Lys Phe Ile Lys	
1 5	
GTC ATG ATG ATC CTC TTC AAT CTA CTC ATC TTT CTC TGT GGT GCA GCC	221
Val Met Met Ile Leu Phe Asn Leu Leu Ile Phe Leu Cys Gly Ala Ala	
10 15 20	
CTG TTG GCT GTG GGA ATC TGG GTG TCC GTC GAC GGG ACA TCT TTC CTG	269
Leu Leu Ala Val Gly Ile Trp Val Ser Val Asp Gly Thr Ser Phe Leu	
25 30 35 40	
AAG GCC TTC GGA TCA CTA TCA TCG AGT GCC ATG CAG TTT GTC AAC GTG	317
Lys Ala Phe Gly Ser Leu Ser Ser Ser Ala Met Gln Phe Val Asn Val	
45 50 55	
GGC TAC TTC CTC ATC GCC GCT GGT GCT GTA CTC TTT ATT CTT GGT TTC	365
Gly Tyr Phe Leu Ile Ala Ala Gly Ala Val Leu Phe Ile Leu Gly Phe	
60 65 70	
CTG GGC TGC TAT GGT GCT CAC TCC GAG AAC AAG TGT GTG CTC ATG ATG	413
Leu Gly Cys Tyr Gly Ala His Ser Glu Asn Lys Cys Val Leu Met Met	
75 80 85	
GTG TGT TGG ACC CAG CTC CCC AGG CCC ACA ACC AAG GGT TCC CCC TTC	461
Val Cys Trp Thr Gln Leu Pro Arg Pro Thr Thr Lys Gly Ser Pro Phe	
90 95 100	
ATC CCT GAC ACC AGC CTC TGG GAT CCC AAA ACC CTA CTC TGG ATT CTG	509
Ile Pro Asp Thr Ser Leu Trp Asp Pro Lys Thr Leu Leu Trp Ile Leu	
105 110 115 120	
GAA GCC TCG CTG TGG GGA CCC CCT GCC TCT TAT GGG AGT GTG CAG GAA	557
Glu Ala Ser Leu Trp Gly Pro Pro Ala Ser Tyr Gly Ser Val Gln Glu	
125 130 135	
GGT GTG TAT CCC TAAAACGTGT TTCCTTTCTC CACTCAGTTC TTTCCATTC	609
Gly Val Tyr Pro	
140	
TCCTCATCAT CTTCATTGCG GAGATTGCAG GGGCCGTGGT TGCTTTGGTG TACACCACAA	669
TGGTGAGGCA CGGGATGGGG CATGGGAGGA AACTGGAGC AAAGCTCTGA GAGTTGGTGG	729

TGGCCTGTGA ATGGCTACTT CCCCTACCAG TCTATTAATA CTGGCCTGAC CCGTCCAGGC 789  
 TGAACAATTC CTGACATTTT TGGTGGTGCC TGCTATCGAA AAGGACTATG GTTACCAGAC 849  
 TGAATTCACC CAAGTGTGGA ACTCTACGAT GGAGGGGGTA AGGTGGGCCG GGAGGATTTG 909  
 GGAATGGAGA GGAGAAAAAA ATCAAGGCAC CATCAGTTAC CAACCCCATT TTGTCTGCAG 969  
 CTGCATTGCT GTGGCTTCAA CAATTACACA GATTTTAACT CCTCACGTTT CGTCAAAGAG 1029  
 AACAAAGTCT TTCCTCCGTT CTGTTGCGCC AACAACTG ACAGTCATAC AGTTGAACCA 1089  
 TGCACCGAGG ACAAGGCCAA AAGTATGAAC GTACAGGTAT GTGCTGGTGT GGAATTAGCT 1149  
 TTTTCGTTT TGATCCGTT TCTCTGTGTA GCTGTGGATG TCCTGGAAGT CGATATGTAG 1209  
 ACCAGGCTGG GCTTGAACAG AGATCGACCT GCCTCTGCCT CCTGAGTGCT GGGGTAAAG 1269  
 GCCACCATCA CCACCCAGCG CCAGTGTAAT TGCTTTAATG GCTTTACAGT AAGGAGTAAA 1329  
 GGGGTGGAG GTAGGGCACT GCCAGCTATA AATACTTTCT TCTCTCTCCC AAGGGTTGTT 1389  
 TCAAACAGAT TCTGCAGAAA ATCAGAACCA ATGCAGTTAC CGTGGGTGGT GTGGCAGTTG 1449  
 GAGTTGCGGC CCTAGAGCTG GCTGCCATGG TGGTATCCAT GTATCTATAC TGCAATCTGA 1509  
 AGTAAGACTG CTTCTTCCTC CTGACTTGTC GTCACATGGA AACCAGGAAG AGGCACTTGG 1569  
 CACAGCTAAG CAGCAGCAAT GGGTGGAGAT GAAACGTCAT TTGGGCTGGA GTGGATCTGC 1629  
 CTTTGCTGCT CTGGACTTCG GGCTAATTAG GGACCACCGC TCGGGCTGTG TCTGTCTTGT 1689  
 TCTTCCACTA CTGGGGCGAG GGGGGCATTG GTCCCAGTCT CTGAGGTAGC CATGTCTGCC 1749  
 CATCTCCTGA GATTCTTGAT GATCCCAGCA CTCCACTGGT GGATGAAGGC ACTTATAGCC 1809  
 TGGGCATATT CTAGCACTGC AAAGTCATAG AAGGAATGTA CAGAAGGTGT TTTGGAAGTT 1869  
 ATAAACCAAT GAAAATAAAA A 1890

## (2) INFORMATION FOR SEQ ID NO:73:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 140 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Met Gln Cys Phe Lys Phe Ile Lys Val Met Met Ile Leu Phe Asn Leu  
 1 5 10 15  
 Leu Ile Phe Leu Cys Gly Ala Ala Leu Leu Ala Val Gly Ile Trp Val  
 20 25 30  
 Ser Val Asp Gly Thr Ser Phe Leu Lys Ala Phe Gly Ser Leu Ser Ser  
 35 40 45  
 Ser Ala Met Gln Phe Val Asn Val Gly Tyr Phe Leu Ile Ala Ala Gly  
 50 55 60

Ala Val Leu Phe Ile Leu Gly Phe Leu Gly Cys Tyr Gly Ala His Ser  
 65 70 75 80  
 Glu Asn Lys Cys Val Leu Met Met Val Cys Trp Thr Gln Leu Pro Arg  
 85 90 95  
 Pro Thr Thr Lys Gly Ser Pro Phe Ile Pro Asp Thr Ser Leu Trp Asp  
 100 105 110  
 Pro Lys Thr Leu Leu Trp Ile Leu Glu Ala Ser Leu Trp Gly Pro Pro  
 115 120 125  
 Ala Ser Tyr Gly Ser Val Gln Glu Gly Val Tyr Pro  
 130 135 140

## (2) INFORMATION FOR SEQ ID NO:74:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2181 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW083

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (1950..2181)
- (D) OTHER INFORMATION: /label= SAC\_24528

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

GTCGACCCAC AATCGCCCGA GCACTTGGA	TCACAGCAGA AACGACCCAA TCCTGAGCTC	60
CGTAGAAATG TGACCATCAA AACTGAACTC	CCACTAGGCG CACGGCGAAA GATGAAGCCA	120
CTGCTCCAC GGGTTAGCTC ATACCTGGTG	CCCATCCAGT TCCCGGTGAA CCAGTCCCTG	180
GTGTTACAGC CCTCGGTGAA GGTTCCTTG	CCTCTGGCAG CATCTCTTAT GAGCTCAGAG	240
CTTGCCCGTC ATAGCAAGCG AGTCCGCATT	GCACCCAAGG TGCTGCTATC CAACGAAGGG	300
ATAGCCCCAC TTCCTGCCAC AGAACCCATG	AAGGAGGAGA AACCCCTGCT TGGAGAAGGG	360
CTATTGCCTT TGCTTCCTAT TCAGTCCATT	AAGGAAGAAG TAATTCAGCC TGGGGAGGAC	420
ATACCACACT TAGAGAGGCC TATCAAAGTG	GAGAGCCCTC CCTTGGAAGA GTGGCCCTCT	480
CCGTGTGCAT CAGTGAAAGA GGAAGTGTCC	AACTCCTGGG AAGATTCTTC CTGCTCTCCT	540
ACCCCAAAGC CCAAGAAGTC CTATTGTGGG	CTTAAGTCCC CAACACGGTG TGTCTCAGAA	600
ATGCTGGTGA CAAAGCGGAG AGAGAAGAGA	GAGGTGAGCC GATCTCGGAG GAAGCAGCAC	660



```

CTTCAGCCAC CCTGTCTAGA TGAGCCTGAA CTCTTCTTCT CAGAGGACTC CAGCACATTT 720
CGGCCAGCCA TGGAGATCCT GGCAGAGTCT TCAGAGCCTG CACCACAGCT CAGCTGCCCT 780
CAGGAGGAGG GAGGGCCCTT CAAGACCCCC ATCAAGGAGA CATTGCCTGT CTCCTCCACT 840
CCTAGCAAGT CTGTGCTCTC TAGAGACCCT GAGTCCTGGA GGCTCACACC CCCAGCCAAA 900
GTTGGGGGGT TAGATTTTCTG CCCAGTACGA ACCCCCCAGG GTGCCTTTGG CCCTCTGCCT 960
GACTCGCTGG GGCTTATGGA GCTGAATACC ACACCTCTGA AAAGTGTTCC CCTCTTCGAC 1020
TCACCCCGGG AGCTCCTTAA CTCAGAAGCC TTTGACCTTG CCTCTGATCC CTTTAGCAGT 1080
TCTCCACCAC CACATTTGGA AGCCAAGCCA GGCTCCCCCG AGCTGCAGGT CCCCAGCCTT 1140
TCAGCCAACC GTTCTCTCAC AGAAGGCCTT GTCCTGGACA CAATGAATGA TAGCCTCAGC 1200
AAGATCCTTC TAGACATCAG TTTCCCTGGC CTGGAGGAGG ACCCTCTGGG CCCTGACAAC 1260
ATCAACTGGT CTCAGTTCAT CCCTGAGCTG CGATAGAGGC AGGGTCTTAC CCTTGCCACT 1320
CAAGCCACCA GTTATCCTGG CACTTGTGTG GCTGGATAGT GCAAGGCTCA GTGTACCCCA 1380
AACCGTCTGA GGGAGCTAGC AGGCAAGGGC TGAGCGGTGC CTTTGGACCT AATTATGCCA 1440
AGGTAAAAGC CACGCTAAG CCACTGCTGG GACCTATGCA AGCAATAGGA TCTCCAGAG 1500
TCCTCCACTC CCTGCTGGCA AGTGAAGTGG GTGTGACAGA GCCGTGAGGA CCAGGAAATG 1560
CCCACCCATT AGTCACCTGC TGCTCCTGGC AGGATAACCC TTGTAAATGG TGTCAGTTCC 1620
CCAAGTTGTC CTGTAATTAT AAATGTAGCC ATATTCCCTT AGCTCTCATT ATCCAGAGAC 1680
TGCCAGGATG GGTAGGGTGA CAAGGGGTTG CATTAGCTTC TGCTTGTGGC CTTTGGGGGC 1740
AGGACCTGCA GTTCAGCCTC TTCACACTGT GGGTTCTGCT GTAGGCTTCT AGACACACAG 1800
GTGTCCTTGC CAGGACCCCA CTTACTGCCC TTTCTCACA GCTCCCCCTG GTTCTAAGCC 1860
AGTGGTACTG CATGAAGAAA TCCTGCGGCA AAGCCTATTG GCTCTGGGTG TGTGGGGACG 1920
GGTGTGCCTG AAGCAAAAGC ATGGGTACTC ACGTGAGTCC TTTAGGTGTT TCTCTGATCG 1980
TGTTCCCAAT CATGCCAGGG AGTCTAGCAT TGAGAACTCA GGCTGAGGCC TGAGGAGGAG 2040
GAGGAAGTGA CCACTGACTT GCCTGGCTTC CTTAGCTTGC ACCTGAGTTT TGCAAAAAGC 2100
CACCCTAGAC CCCACTCTAC AAGCTAGCAC AAGAACACTA CTGTAACCTAC CTACTGAATA 2160
AAGCCCAGGT GGCCTGAAAA A 2181

```

## (2) INFORMATION FOR SEQ ID NO:75:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3336 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW084

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (163..391)  
(D) OTHER INFORMATION: /label= SAC\_24854

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

TTTTTCCACC AAGAGGGAGT CCTTTAATAG ATTGCACAGA CAGACACAAC AGGCCAGCAC	60
CCTGCTTGCA GGCCAGGCG CTTTGCCAGC AGATGCAATG GCCTCCCCCA CCTGGGTGTG	120
GGCCGAGGG TGGAGATGAG GTGGTTTAAA AAACACAGCT GTACTAATTC TTCACTTTCA	180
CAGAAAAGAG AAAC TTGGGG GAGCAGCTCA CAGGAGAATT TACATAGAGC TGAGTCTGAA	240
TGTCCACCCC TAGAGAAGTT GAACCTCCCC CTGAGCTGTC TGTGCTCAGC AGGGGGCAGA	300
CAGCACATGG AGGGCGCAGG CCAGTCCACA GAGGAGCGCT GACTTAAACA GCTGTCTCC	360
CTTCCCAGNT GTTAGNATTT CTACACNATG TACATTGTTG TGCTTTATTG CTCCCACCCC	420
AGCCTCTGCC TGAAGCCCAG GGTGGAGGGA GGAACACCTG AACCAGATTC AGCTGGGTGG	480
GAGCACCAAG CCCCCAATGG CTGGTGCTCT GTAAGNTTCT GAGGCTGCCA GGCTCNNGNC	540
ACTCGGAGCT GCAGACCAAG GGTACAGAGC CAGCCGACCA GCCAGCTCTT CTGAGGAGCA	600
GCAGAAACAG AAGGATCAGT GCTGGACAAG ATGCAGGCTG GGTAGGAGCA GGCAGGGTGG	660
AAGGACGAGT GCAAGGCCAT GAGTCTGGGG CAGAGGGTGG TGCACATGGG GCTGTGGGGG	720
CCTGCAGGGC ACCTCACGAG ACAGAACCAT TCTGTAGCAG AGCTTGACGG AGGAAGGGTG	780
TGAGGACCAC TGGCCTGGGG CCCCTTGGAG CACAGAGCTG CTTCCAGTTC ATCAGGGGCT	840
GGCAGTGCCA GGCCCCAGCT CGGGCGGTGT GCATTGGGAG ATTCAGGGAG AACCNGCAGG	900
CAGCCTGTGG GGCTGGGTGA GGAGGGAAGG CTGTGGTGGT ATCTTTGAGT NTGTCTTGGC	960
CGGGCCAGGG GTTAACACAG GATCAGTTGT TCTCAAAGAG TGAGAGCAGA TCATCACTGC	1020
TGTTTGTGGG GAGGTCAGGT GGACCCAGGT AGGAGAGTAG TTCATCAGGA TTGGTCAGTT	1080
CTGGGAGCAG ANCAGAGCT GGTCTGAGG CCTCCCCTGC CCCAGTCATG CTGGGAGGCC	1140
CCATCATGCC CGTGGCAGNA TTGAAAGCCA GTTCACTGAT GGGCCCTGTG TTTGGTTGGC	1200
CCAGGGGCTG CCTGGATGCA GGGGAAGGGT TTGGATGGTG TAGCTGGGTG CCTGGGGTGG	1260
GCCCCAGGTT GGGAGTGTGC AGTCCTGGTG GTCCAGGATT ATGGGCTGGG TCCCAGGTGA	1320
CCCTGCCTGG TGCCATCCTT GGCTGGGAG GCAAGGAGCA GACTTGNNNN GAGTCAGGAG	1380
GCTGCTGGGA ATGTCAGACT GGTAGGAGAT TGGGGGTGGC CCCTGGGGTG AACTCAGCAA	1440
GGTTTGGGGT GSCGGGTGTA GCAGATGGGA AAGATTCTGG GAAGGTTCCA GGGCCCCAGG	1500

AAGTTGGAAC CCTGGCTGGG NTAGTCGNTG GGGGTAGGGN CCNAAGGGGC TGCAATGGGG	1560
CAAANGGGCA GCGCCAGGGC CCAGGGCCGC AATCATCTCC ATCACACTGG GCATGAGCAC	1620
GTGGGCAGGG CTCACAGTAC GGCAGCGCTT CAGCACTGGG CCCGTCCGGC TCCTCCTTGA	1680
TGTGTAGGTC AGGCTTCACA GGCAC TGCTT TCCAGCTGCA TGTGGGGTCG ATGGTGATCT	1740
CCTCGTAGTC TGAGTTCTGA ATATAAATCA GGATGCCTAG CATATACTGA TCCACCTCCA	1800
GGCCCTCCAG CAATGCCGTC TTGTTGCACA CCGGGACACC TCCAGGTCCC CCGCTCACAG	1860
TTAAGCTGTA AGTATGACTC CAGGTCAAAG CACTGTATGT GGCGACAGTC GTGACCACGA	1920
GCTGGGAGCT GGATCCTGCG GAAGGTGATG GGGCACTTCA GGGACACCTT GATAGCCGTC	1980
TGCTCCACCC CATCCTCTCC ATTGGGCCCCA GGGGTGCCAG GGATGGTGCC GCTACTGAAG	2040
TTCCGCTTTA TCTTGGTGAT GCAGTGCTCA GCTGGCAAGN GGCCTTCTT GAGGAGACCC	2100
TGCAGCACAG AGCGGACAGA CGGACGGTGC ACCAGCTGCA GCACGAAGAG GTGGGAGCAG	2160
CAGCAGGCGG TGACGGTGAT CTGGATGGTG TTGCGACCCG GCTGGCACAC ATGCTTCAGG	2220
TAGAGGGGCT TGTGCGAGGT CTTGTTGTCT CCACGCTCGA TGCTGAGGGG TGTGNCATTG	2280
ACGCTGACCT GCACCGNNGC CGGCCANTTG GTGGTTCATC TGCCGGTCCT CGTGGTGGTA	2340
GCACTTGAAC TCGGAAGTGA GCTCTAAGTC AGGTCTCAGC ATCAGGGTTT TATACACAGA	2400
GTCTCGNAGC TGGAAGACAT GGTGCTCAC AGCCAGGTTG TGCTGCAGGC GGAAGGGCTC	2460
CAGGACCACC CCATCCCGAA CTGGGAAGGT CAGGCGCAGC TCGTCACAGG GGACGCTTCC	2520
AGAGGGTGAT GGGTGCAAGG AGGCTGAGGC CTGGCTTGAG GTCAGGCAGG AAAGGAGACT	2580
TGACCTCCTG GCTTGGGGAC ACGTAGGGAG CACTGCTGCT AGGCGTCATG GGTGGTGTGG	2640
GATTCCCCGG CAGTGGGGAG CTGGGGTAAG CAGGGATGGA GCGGGAAGGC CCACTCAAGC	2700
CAGGCTGGCT GTAGCTGATG CTGCCCCCGT TGAANCTGGC GCCCTGCCCG TTGAACTGCT	2760
CTGTGGGCTT GTAATGCAGN CCTGTGGGGC CAGGGGCAGA CAGGGACTGG GCCATGCCCT	2820
GCTGCAGGGG GAGTCTGTGT CCTGCGTAGG AGGAGGCAGG GCCAGGGGGC TGCCCCGGAC	2880
CAGGAGCATA CTGGGCAGTG CTGGCTGCAT ACTGGCCNNC TTGCAGATAT TGNTGNCNAG	2940
GATACACCTC TGAGTAGGTC TCTCTTGACC CNCTGTGAG GCAGCGGCTG CCCCTGGGGA	3000
GGCCCAGGGT AGCCATGCTG AGGCAGTCGC TGTCTGTCAT ACAAGGGTGT CATGCCTGCT	3060
GCCCCTGTAG GATTCATGGC CATGGAGGAA AGGCCTGAGG GGCCCATCAT TCCTCCCATG	3120
CCTGAAGGGT TCATGCCAGG GGGTACACTG GGCCCTCGAG GACCTCCATG CTGCAGAAAC	3180
TGGCTGTAA AAGATTGCCC AGTCCCCATC GCTCCATATT GGCTTAGCTT CCTGGCTCTN	3240
CTTCTCTTGC AGGNNGCCA CAGTAGCTGT GGCTGTGGCT GTGGCCGTNC TGCAGCAGCA	3300
GCCATGGCAG CTGCAGCTCT GCTTGTGTGA AGTCAG	3336

## (2) INFORMATION FOR SEQ ID NO:76:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1494 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW088

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 9..686

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (612..909)
- (D) OTHER INFORMATION: /label= SAC\_24336

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GTGACCC GCT GCG GTC TAT CTG GGC AGC ACC ATG CTC GCG CTC CGA AGC	50
Ala Ala Val Tyr Leu Gly Ser Thr Met Leu Ala Leu Arg Ser	
1 5 10	
GGC CTG AGG ACG GCG CTG GCC CCG CGG GCT CTG ACG CCT CAG GTA TGT	98
Gly Leu Arg Thr Ala Leu Ala Pro Arg Ala Leu Thr Pro Gln Val Cys	
15 20 25 30	
TCA CCT TTT GCT ACA GGC CCC AGG CAA AGC GAT GGG ACA TTG TAT GAA	146
Ser Pro Phe Ala Thr Gly Pro Arg Gln Ser Asp Gly Thr Leu Tyr Glu	
35 40 45	
TTC CGC ACC TAC TCT CTG AAG CCC TCA AAG ACA AAT GCG TTC CTG CAG	194
Phe Arg Thr Tyr Ser Leu Lys Pro Ser Lys Thr Asn Ala Phe Leu Gln	
50 55 60	
AAT TTT CAG AAA TAC GTT CAC CTT CGG ACA GCT CAC TCT GAG ATG ATT	242
Asn Phe Gln Lys Tyr Val His Leu Arg Thr Ala His Ser Glu Met Ile	
65 70 75	
GGA TAC TGG ACC GTA GAA TTT GGA GGC AAA ATA AAC AGA GTG TTC CAT	290
Gly Tyr Trp Thr Val Glu Phe Gly Gly Lys Ile Asn Arg Val Phe His	
80 85 90	
ATT TGG AAG TAT GAT AAT TTT GCT CAT CGA ACT GCC GTC CGC AAA GCC	338
Ile Trp Lys Tyr Asp Asn Phe Ala His Arg Thr Ala Val Arg Lys Ala	
95 100 105 110	
TTG GCC AAG GAT AAG GAA TGG CAA GAA CAA TTC CTC ATT TCA AAT TTG	386
Leu Ala Lys Asp Lys Glu Trp Gln Glu Gln Phe Leu Ile Ser Asn Leu	
115 120 125	
GCT TTC ATG GAT GAA CAA GAG GTT GAG ATT ACC TAC TTG GTA CCC TGG	434
Ala Phe Met Asp Glu Gln Glu Val Glu Ile Thr Tyr Leu Val Pro Trp	
130 135 140	

TGT AAA ATA AGA ACA CCT CCC AAG GAA GGA GTC TAT GAA CTG GCT ACC Cys Lys Ile Arg Thr Pro Pro Lys Glu Gly Val Tyr Glu Leu Ala Thr 145 150 155	482
TTT CAG ATG AAG CCT GGT GGC CCA GCT CTA TGG GGT GAT GCA TTC AAA Phe Gln Met Lys Pro Gly Gly Pro Ala Leu Trp Gly Asp Ala Phe Lys 160 165 170	530
AGG GCA ATA AAC GCC CAT GTT GAG CTT GGC TAC TCT ACA CTA GTT GGT Arg Ala Ile Asn Ala His Val Glu Leu Gly Tyr Ser Thr Leu Val Gly 175 180 185 190	578
GTT TTC CAC ACT GAA TAT GGA GCC CTC AAC AGA GTT CGG GAA AGT GTC Val Phe His Thr Glu Tyr Gly Ala Leu Asn Arg Val Arg Glu Ser Val 195 200 205	626
AAT TAC CTC GAG TCT CAG CAG AAT ATG TTC CTG ATT CCA ACC TCA TTT Asn Tyr Leu Glu Ser Gln Gln Asn Met Phe Leu Ile Pro Thr Ser Phe 210 215 220	674
TCA CCA TTG AAG TAGTTTCCTA CCGAGCGCTG AGCATTCCCA TCACGGAGGT Ser Pro Leu Lys 225	726
GAGATGCGTT TCCCCCAGAC AATCGCCCTG TTCATTGTCT GAAGGAGGTG CTGAGCTACC	786
GCACTGCCCT CTGCATGTCT GAGACTTCAT GTATCCTTTA CTACCAGTCG CTCAGGAAAT	846
GTTTCTGTCC CTTTTCATGT CTACTGGATG TCAGAAAGTG GTGTTATTAC TCCATGCCCT	906
GGTACTTCAT TGTTCTTCTG TGGTTTTATG ATTAAGTCTT GTCCTCAAAA TATGATTTTG	966
AAATTTTTTC CCTCAAAATA TTTGAAATT AAATAAATAC AATATCTGCC TTTTATTGTC	1026
GTTCATTTTC CCTGAAATTT TTAAGTGTGTT TTACCAGTAA TGTTTACTTT GTGATTTCAA	1086
ATGGGATTGA TAGAAGTCAG AGAAAGAACT GGAAGAGCTT GAAGGGGCTC GAGACCCCAT	1146
ATGTACAACA ATGCCAAGCA ACCAGAGCTT CCAGGGACTA AGCCACTACC TAAAGACTAT	1206
ACATGGACTG ACCCTGGACT CTGACCTCAT AGGTAGCAAT GAATATCCTA GTAAGAGCAC	1266
CAGTGAAGG GGAAGCCCTG GGTCTTGCTA AGACTGAACC CCAGTGAACG TGATTGTTGT	1326
GGGGAGAGTG GTAATGGGGG GAGGATGGGG AGGGGAAGCC CATATAGAAG GGGAGGGAGA	1386
GGGGTTAGGG GGATGTCGGC CCAGAAACCA GGAAGGGGAA TAACAATTGA AATGTAAATA	1446
AGAAATACTC AAGTTAATAA AGATGGAGAA AAGAAAAAGA AAGTCGAC	1494

## (2) INFORMATION FOR SEQ ID NO:77:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 226 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Ala Ala Val Tyr Leu Gly Ser Thr Met Leu Ala Leu Arg Ser Gly Leu  
 1 5 10 15  
 Arg Thr Ala Leu Ala Pro Arg Ala Leu Thr Pro Gln Val Cys Ser Pro  
 20 25 30  
 Phe Ala Thr Gly Pro Arg Gln Ser Asp Gly Thr Leu Tyr Glu Phe Arg  
 35 40 45  
 Thr Tyr Ser Leu Lys Pro Ser Lys Thr Asn Ala Phe Leu Gln Asn Phe  
 50 55 60  
 Gln Lys Tyr Val His Leu Arg Thr Ala His Ser Glu Met Ile Gly Tyr  
 65 70 75 80  
 Trp Thr Val Glu Phe Gly Gly Lys Ile Asn Arg Val Phe His Ile Trp  
 85 90 95  
 Lys Tyr Asp Asn Phe Ala His Arg Thr Ala Val Arg Lys Ala Leu Ala  
 100 105 110  
 Lys Asp Lys Glu Trp Gln Glu Gln Phe Leu Ile Ser Asn Leu Ala Phe  
 115 120 125  
 Met Asp Glu Gln Glu Val Glu Ile Thr Tyr Leu Val Pro Trp Cys Lys  
 130 135 140  
 Ile Arg Thr Pro Pro Lys Glu Gly Val Tyr Glu Leu Ala Thr Phe Gln  
 145 150 155 160  
 Met Lys Pro Gly Gly Pro Ala Leu Trp Gly Asp Ala Phe Lys Arg Ala  
 165 170 175  
 Ile Asn Ala His Val Glu Leu Gly Tyr Ser Thr Leu Val Gly Val Phe  
 180 185 190  
 His Thr Glu Tyr Gly Ala Leu Asn Arg Val Arg Glu Ser Val Asn Tyr  
 195 200 205  
 Leu Glu Ser Gln Gln Asn Met Phe Leu Ile Pro Thr Ser Phe Ser Pro  
 210 215 220  
 Leu Lys  
 225

## (2) INFORMATION FOR SEQ ID NO:78:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1082 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (iii) HYPOTHETICAL: NO

## (vii) IMMEDIATE SOURCE:

- (B) CLONE: HW089

## (ix) FEATURE:

- (A) NAME/KEY: CDS

(B) LOCATION: 8..1006

## (ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 961..1076

(D) OTHER INFORMATION: /label= SAC\_24461

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

GTCGACC	AAG	CAA	CAG	CAG	GGA	AGT	GAA	TTT	CAG	ACA	CTG	CTG	GAG	GCT	49	
	Lys	Gln	Gln	Gln	Gly	Ser	Glu	Phe	Gln	Thr	Leu	Leu	Glu	Ala		
	1				5					10						
GAG	GAG	ACC	AAA	GCC	GCC	TCT	GAC	GCT	TCA	AAT	AAA	AAG	CTT	TGG	TCC	97
Glu	Glu	Thr	Lys	Ala	Ala	Ser	Asp	Ala	Ser	Asn	Lys	Lys	Leu	Trp	Ser	
15				20					25					30		
CTC	AGA	ACT	CTA	CCC	TTC	CAG	CCC	TGC	TCT	GAA	GCC	TCA	GAC	GTG	AGA	145
Leu	Arg	Thr	Leu	Pro	Phe	Gln	Pro	Cys	Ser	Glu	Ala	Ser	Asp	Val	Arg	
			35					40					45			
ACC	ATG	TCT	GTG	ACC	CTG	CAA	GCT	GTC	CTC	TAC	AGT	CTT	CAG	GAG	GAG	193
Thr	Met	Ser	Val	Thr	Leu	Gln	Ala	Val	Leu	Tyr	Ser	Leu	Gln	Glu	Glu	
	50						55					60				
CAG	TCC	AGG	CTC	AAG	ACG	AGG	CTG	CAG	GAA	CTA	CAG	CAG	CTG	AAA	AGG	241
Gln	Ser	Arg	Leu	Lys	Thr	Arg	Leu	Gln	Glu	Leu	Gln	Gln	Leu	Lys	Arg	
	65					70					75					
GAG	CGC	ATA	GAT	GTT	CCC	AGA	GAC	AAG	ATC	CCA	TTC	CCG	GTG	CCT	GAA	289
Glu	Arg	Ile	Asp	Val	Pro	Arg	Asp	Lys	Ile	Pro	Phe	Pro	Val	Pro	Glu	
	80					85					90					
GTC	CCC	GTG	GTA	TTC	CAA	GGT	CTA	ACT	AAG	CAG	AGC	GGC	CAA	GCA	CCC	337
Val	Pro	Val	Val	Phe	Gln	Gly	Leu	Thr	Lys	Gln	Ser	Gly	Gln	Ala	Pro	
95					100					105					110	
AAG	TCT	CTA	GTT	TCT	AAC	TTG	AAG	ATC	TGC	TGC	CCT	CTG	CCT	GGA	GGT	385
Lys	Ser	Leu	Val	Ser	Asn	Leu	Lys	Ile	Cys	Cys	Pro	Leu	Pro	Gly	Gly	
			115					120						125		
TCT	GCT	CTG	ATC	ACC	TTT	GAT	GAC	CCC	AAA	GTG	GTT	GAG	CGG	TTG	CTA	433
Ser	Ala	Leu	Ile	Thr	Phe	Asp	Asp	Pro	Lys	Val	Val	Glu	Arg	Leu	Leu	
		130						135				140				
CAA	CAA	AAG	GAA	CAT	AAA	GTT	GAC	ATC	GAG	GAC	TGC	CGG	CTG	CGG	GTG	481
Gln	Gln	Lys	Glu	His	Lys	Val	Asp	Ile	Glu	Asp	Cys	Arg	Leu	Arg	Val	
	145					150					155					
CAG	ATC	CAG	CCC	TTG	GAG	CTG	CCT	ATG	GTG	ACC	AAC	ATC	CAG	GTG	TCC	529
Gln	Ile	Gln	Pro	Leu	Glu	Leu	Pro	Met	Val	Thr	Asn	Ile	Gln	Val	Ser	
	160					165				170						
AGC	CAG	CCG	GAT	AGC	CAC	AGG	GTG	CTA	GTC	AGT	GGC	TTT	CCT	ACT	GGA	577
Ser	Gln	Pro	Asp	Ser	His	Arg	Val	Leu	Val	Ser	Gly	Phe	Pro	Thr	Gly	
175				180					185					190		
CTT	AGG	CTG	AGT	GAG	GAG	GAG	CTA	CTG	GAC	AAG	TTA	GAG	ATC	TTC	TTC	625
Leu	Arg	Leu	Ser	Glu	Glu	Glu	Leu	Leu	Asp	Lys	Leu	Glu	Ile	Phe	Phe	
		195						200				205				
GGC	AAG	GCC	AAG	AAT	GGA	GGT	GGG	GAT	GTG	GAG	ACC	CGG	GAG	ATG	CTG	673

Gly Lys Ala Lys Asn Gly Gly Gly Asp Val Glu Thr Arg Glu Met Leu	
210 215 220	
CAA GGG ACT GTC ATG CTG GGA TTT GCT AAT GAA GAA GTG GCC CAG CAC	721
Gln Gly Thr Val Met Leu Gly Phe Ala Asn Glu Glu Val Ala Gln His	
225 230 235	
TTA TGC CAG ATT GGC CAG TTC AGA GTC CCA CTG GGT CGG CAG CAG GTC	769
Leu Cys Gln Ile Gly Gln Phe Arg Val Pro Leu Gly Arg Gln Gln Val	
240 245 250	
CTT TTG AGG GTC TCT CCC TAT GTG AGT GGG GAG ATC CAG GAA GCT GAG	817
Leu Leu Arg Val Ser Pro Tyr Val Ser Gly Glu Ile Gln Glu Ala Glu	
255 260 265 270	
ATC AAA TTC CAG CAA GCA CCT CAT TCA GTG CTG GTG ACG AAT ATT CCC	865
Ile Lys Phe Gln Gln Ala Pro His Ser Val Leu Val Thr Asn Ile Pro	
275 280 285	
GAC GTC CTG GAT GTG CAG GAA CTG CAT GAC ATT CTG GAG ATC CAC TTC	913
Asp Val Leu Asp Val Gln Glu Leu His Asp Ile Leu Glu Ile His Phe	
290 295 300	
CAG AAG CCT ACC CGT GGG GGC GGA GAG GTG GAG GCC CTG GCA GTT GTA	961
Gln Lys Pro Thr Arg Gly Gly Glu Val Glu Ala Leu Ala Val Val	
305 310 315	
CCT GTG GGG CAG CAG GGC CTG GCT GTC TTT ACT TCT GAG TCA AGC	1006
Pro Val Gly Gln Gln Gly Leu Ala Val Phe Thr Ser Glu Ser Ser	
320 325 330	
TAGGGCTGCA TGTCCACCTG CTGGGCATGG GACTGCTCAT TGCCTAATGC TACCAAGAAC	1066
AGTGAAAGGC GTCGAC	1082

## (2) INFORMATION FOR SEQ ID NO:79:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 333 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Lys Gln Gln Gln Gly Ser Glu Phe Gln Thr Leu Leu Glu Ala Glu Glu	
1 5 10 15	
Thr Lys Ala Ala Ser Asp Ala Ser Asn Lys Lys Leu Trp Ser Leu Arg	
20 25 30	
Thr Leu Pro Phe Gln Pro Cys Ser Glu Ala Ser Asp Val Arg Thr Met	
35 40 45	
Ser Val Thr Leu Gln Ala Val Leu Tyr Ser Leu Gln Glu Glu Gln Ser	
50 55 60	
Arg Leu Lys Thr Arg Leu Gln Glu Leu Gln Gln Leu Lys Arg Glu Arg	
65 70 75 80	
Ile Asp Val Pro Arg Asp Lys Ile Pro Phe Pro Val Pro Glu Val Pro	



85 90 95

Val Val Phe Gln Gly Leu Thr Lys Gln Ser Gly Gln Ala Pro Lys Ser  
100 105 110

Leu Val Ser Asn Leu Lys Ile Cys Cys Pro Leu Pro Gly Gly Ser Ala  
115 120 125

Leu Ile Thr Phe Asp Asp Pro Lys Val Val Glu Arg Leu Leu Gln Gln  
130 135 140

Lys Glu His Lys Val Asp Ile Glu Asp Cys Arg Leu Arg Val Gln Ile  
145 150 155 160

Gln Pro Leu Glu Leu Pro Met Val Thr Asn Ile Gln Val Ser Ser Gln  
165 170 175

Pro Asp Ser His Arg Val Leu Val Ser Gly Phe Pro Thr Gly Leu Arg  
180 185 190

Leu Ser Glu Glu Glu Leu Leu Asp Lys Leu Glu Ile Phe Phe Gly Lys  
195 200 205

Ala Lys Asn Gly Gly Gly Asp Val Glu Thr Arg Glu Met Leu Gln Gly  
210 215 220

Thr Val Met Leu Gly Phe Ala Asn Glu Glu Val Ala Gln His Leu Cys  
225 230 235 240

Gln Ile Gly Gln Phe Arg Val Pro Leu Gly Arg Gln Gln Val Leu Leu  
245 250 255

Arg Val Ser Pro Tyr Val Ser Gly Glu Ile Gln Glu Ala Glu Ile Lys  
260 265 270

Phe Gln Gln Ala Pro His Ser Val Leu Val Thr Asn Ile Pro Asp Val  
275 280 285

Leu Asp Val Gln Glu Leu His Asp Ile Leu Glu Ile His Phe Gln Lys  
290 295 300

Pro Thr Arg Gly Gly Gly Glu Val Glu Ala Leu Ala Val Val Pro Val  
305 310 315 320

Gly Gln Gln Gly Leu Ala Val Phe Thr Ser Glu Ser Ser  
325 330

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2263 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW090

## (ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 335..937

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION: 1355..1434  
(D) OTHER INFORMATION: /label= SAC\_24197

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

```

GTCGACCTTT AAGCTTACCA TCAGAACAAAC AAATCAAAAT GTAAACTTAA AATATAACCA      60
AAAGAGGGAC AGCTCTTTAG GAAAAGGAAA AAACCTTAAA TAGTGAATAA ACAACTACAA      120
CCACTTAACC ATTGTAGGCT TAAAGCAGC CATCAATAAA GAAAGCGTTC AAGCTCAACA      180
TACACACACA CACACACACA CACACACGCT CCCTCCCTCC GGCGCTCTAC TCCGCGCCCG      240
CCCGCCCGCT CGCCCGCCCG CCCGCCTCGC ACGCACCGGC CTCGGCTCCC CGCGCCCGCC      300
GCCGAGCGAA GCTGCTCTGG GATCGGAGCA GCTC ATG GAG AAA GTG CCG GGC      352
                               Met Glu Lys Val Pro Gly
                               1           5

GAC ATG GAG ATA GAG CGC AGG GAG AGG AAC GAG GAG CTG TCC GAG GCG      400
Asp Met Glu Ile Glu Arg Arg Glu Arg Asn Glu Glu Leu Ser Glu Ala
          10           15           20

GAG AGG AAG GCG GTT CAG GCT ACG TGG GCC CGG CTG TAT GCC AAC TGC      448
Glu Arg Lys Ala Val Gln Ala Thr Trp Ala Arg Leu Tyr Ala Asn Cys
          25           30           35

GAG GAC GTG GGG GTG GCC ATC CTG GTG AGG TTC TTT GTG AAC TTC CCG      496
Glu Asp Val Gly Val Ala Ile Leu Val Arg Phe Phe Val Asn Phe Pro
          40           45           50

TCG GCC AAG CAG TAC TTC AGC CAG TTT AAA CAC ATG GAG GAT CCG TTG      544
Ser Ala Lys Gln Tyr Phe Ser Gln Phe Lys His Met Glu Asp Pro Leu
          55           60           65           70

GAG ATG GAG AGG AGT CCT CAG CTG CGG AAA CAT GCC TGC CGG GTC ATG      592
Glu Met Glu Arg Ser Pro Gln Leu Arg Lys His Ala Cys Arg Val Met
          75           80           85

GGG GCC CTC AAC ACG GTC GTG GAG AAC CTG CAT GAC CCA GAC AAG GTA      640
Gly Ala Leu Asn Thr Val Val Glu Asn Leu His Asp Pro Asp Lys Val
          90           95           100

TCC TCT GTG CTC GCG CTG GTT GGC AAG GCC CAC GCC CTC AAG CAC AAG      688
Ser Ser Val Leu Ala Leu Val Gly Lys Ala His Ala Leu Lys His Lys
          105           110           115

GTG GAA CCT ATG TAC TTT AAG ATT CTC TCT GGG GTC ATT CTG GAC GTG      736
Val Glu Pro Met Tyr Phe Lys Ile Leu Ser Gly Val Ile Leu Asp Val
          120           125           130

ATC GCC GAG GAG TTT GCC AAT GAC TTC CCC GTC GAG ACG CAG AAA GCC      784
Ile Ala Glu Glu Phe Ala Asn Asp Phe Pro Val Glu Thr Gln Lys Ala
          135           140           145           150

TGG ACC AAG CTG CGC GGT CTC ATC TAC AGC CAT GTG ACC GCG GCC TAC      832

```

Trp Thr Lys Leu Arg Gly Leu Ile Tyr Ser His Val Thr Ala Ala Tyr	
155 160 165	
AAG GAA GTG GGC TGG GTA CAG CAG CCT ACC AGC CAC ACT GCC CTC TTC	880
Lys Glu Val Gly Trp Val Gln Gln Pro Thr Ser His Thr Ala Leu Phe	
170 175 180	
AGG GCC ATA GGA CCC CTC CCA GCT CCC CCC TCC CTG GCA GCA CTT TGG	928
Arg Ala Ile Gly Pro Leu Pro Ala Pro Pro Ser Leu Ala Ala Leu Trp	
185 190 195	
GCA GAA GGC TGACTTCTGA AGACACTCCT TGACCTTCCA TTTCTGGGTG	977
Ala Glu Gly	
200	
CCAAGGAAGC TGGAGGAACC CCTGACTGGA CTTCCCGAAA GGAGGCCTTC ATCATGGCTG	1037
TGGTCTCCCA CGTTACTGGG AGGTGGGGCT GGCTACCTGG ATGCTGAATT TAGCAGAGGG	1097
GAACCCCTTCT GCCTCGCAGC AGAGGGCACA CCCTCCTTAG CTTTCACTGT CAAGAGAGGG	1157
ACCTGGCATG TGGCCGGTGG GAAGCAAGGA CAAGCCTTTG AGTCACTTCA GAAAGGAACT	1217
AGCCCTGCTT CACTCTGCCC TACCAGCATG CAGGTTCTCC TCACCATCTC TGAGCATTGC	1277
GCACACACAC GTCTACCATA TACACAGATA TATTCTATAT ACGAGCTCTA TATAAATATA	1337
TATACACACA CACACGTACA TATCTGTCGC GTGCACCTGC AGGCCAGAA GCCTAGGCCT	1397
CGCACAGCTG TGGGAGACCC GGAAGAGTCC TGGCCGTACT TCAGGGCAGC TCTACCGGGA	1457
GCAGGGAGAA AGCTTGCAAC AGCACGAGAG GCGGGGCCG GGTGGGAGAG TCGGGGTGAC	1517
TCAGCTCACC CCGCACAGCT CAGTGCCCTG CAATCTTCCC CTTCCGCCCC CCAAAGAGA	1577
ACACTTGGTT TGGCCGGTGT CGGGGCGCAC ACCTGTAATC CCCGAGTTTG GGACGCAGAG	1637
GAAGGACAAT CTTTGCAAGT CTGAGGTCAG GTGGTCTACA CTTGGTCTCC ATCTTGGTGA	1697
GTTTCCGGCT AGCTGGGGTC GACTCAGGGA GACCCGTGT CAAACAAAAC AAGAAGCGAA	1757
GAGAAGATCT GGCTCCCCTC CCCTCCCCAA GGGTTTGCCC GACTGACTTG CTCCGGAAAG	1817
GGGCACACCG AGAATTCAGG AGACCCAGT GGGGTGCTGG CCAGTGTAGG TGTCCGGCGT	1877
TTGGGGGTGG GGTGAGGGGC AGCTTCAGCG TGGAGCCCTA GAATGTCAGC CTCACCCGCT	1937
CACTTCAGAC AGCAGCCAGG CCACAGGACC AGGGGCCATA CTCAAACCCC TCAGCAGAGA	1997
CCAGGGCGCC TCTCTCCTGC CCTCCTCGGG GCCACCTGCC CATCCTCTCC AACTAGTCTC	2057
CGTAACATTC GCTGGGGACC CCGTGTGTAC CCCTTGGTCA CGTGTACTGT CCTGGAGGAC	2117
GTCTTACTGC TGTGGCTACC GTGCCTGCGT CCCCTCTGGT CCCACCCTCA CGCAGCCTTC	2177
CTGTAAGTGC CTGTCTTTTT GTAGTTTCTG ATGTTTGTAA CCAGACCCAG CTGTGTCATT	2237
AAACAGGTCT GTTCTTCCTG TAAAAA	2263

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 201 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

```

Met Glu Lys Val Pro Gly Asp Met Glu Ile Glu Arg Arg Glu Arg Asn
 1           5           10           15
Glu Glu Leu Ser Glu Ala Glu Arg Lys Ala Val Gln Ala Thr Trp Ala
      20           25           30
Arg Leu Tyr Ala Asn Cys Glu Asp Val Gly Val Ala Ile Leu Val Arg
      35           40           45
Phe Phe Val Asn Phe Pro Ser Ala Lys Gln Tyr Phe Ser Gln Phe Lys
 50           55           60
His Met Glu Asp Pro Leu Glu Met Glu Arg Ser Pro Gln Leu Arg Lys
 65           70           75           80
His Ala Cys Arg Val Met Gly Ala Leu Asn Thr Val Val Glu Asn Leu
      85           90           95
His Asp Pro Asp Lys Val Ser Ser Val Leu Ala Leu Val Gly Lys Ala
      100           105           110
His Ala Leu Lys His Lys Val Glu Pro Met Tyr Phe Lys Ile Leu Ser
      115           120           125
Gly Val Ile Leu Asp Val Ile Ala Glu Glu Phe Ala Asn Asp Phe Pro
      130           135           140
Val Glu Thr Gln Lys Ala Trp Thr Lys Leu Arg Gly Leu Ile Tyr Ser
      145           150           155           160
His Val Thr Ala Ala Tyr Lys Glu Val Gly Trp Val Gln Gln Pro Thr
      165           170           175
Ser His Thr Ala Leu Phe Arg Ala Ile Gly Pro Leu Pro Ala Pro Pro
      180           185           190
Ser Leu Ala Ala Leu Trp Ala Glu Gly
      195           200

```

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1740 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW092

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
 (B) LOCATION: complement (303..566)  
 (D) OTHER INFORMATION: /label= SAC\_24320

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GTCGACCTCG GAGCTGGGGA CCGAACCCAG GGCCTTGCGC TTGCTAGGCA AGCGCTCTAC	60
CACTGAGCTA AATCCCCAAC CCCTAGCCCT TTGTTTTTTA AGAAAGTCTA GTGTAGCCCA	120
GGTTGGCCTT AAATTCCTGA TCCTCCTGCC CCTATCTGTA ATCTTGGGAT TATAGGTGTG	180
GGCCATCACA CTGTTTATCC AGTGCTGAAT CAAACCCAGA ATCTTCCACA TGCTAGGCCA	240
GCAGTCTCCT AAATACTACA CGGCTATAGG TTCAGCCTAG ATCTCTCTTT TGAAGCTCTC	300
GTACACCTAT AGGGGCCACT AATCCTAAAG ACGCAGAGGA TGAAAACCAC CAGGCCCAAG	360
GGTTAACGTG GAAAACAGAG CCCTGGTGTC CCCCTGGGGA ACCAGCCCTT CATGTTCTAA	420
GACAGTCATC ACCCAGATTC CTGAAAACAA TAGCGGCTTC TTGTTTAGAA CGGAACCGCA	480
CCTGCTGGAA TACGTGTCAA CAAACAATCG CTTTGTTTTT CATATTCGCC ATCCAAGCTG	540
AGGCAATGGG ACTTCCTGAA GCGAGTACTG GAAGACTTAT TCTTAGAAGA GGACAGAATT	600
CTCATTTTAA CACGAGGCTC AGGGGCCGAG TGCTGGCCAT CTTCCCTCCT TCTTTTGTAC	660
TCCCTGGTGT TCATCTTATT TGTTTTTCAA CAGGTTTATT TCCATTGCAG CCCGGGGCTT	720
CCAACGCCCA AAAACGCCCA GCCCAGCATG CCGTAAATGG TGAGCACGCT GGCTTCTGCT	780
TCCTCTCCCT GAGAACCCAG ACCCTTTGTT GTTCCGCCAT TCCAAAAGGC CCTGGGAAGG	840
TTCTTTGCTC CTCTGGTGAC AGGCTTGGGA TGAAGACTTG GGAACCTCTA TTTGCTAAAC	900
TCTCGCTCCT GGCTAAATCT TTCTGCTTGT TCTTTTTCCT CGTCTAATCA TGCTACCTGC	960
TACTACTTTT CCTGATGCAG TCTGACGAAC GTCTGTTGAA CACATCGGGA CAGTTTAAAA	1020
AGGAGTTATT TTGGTGCCAA AGGCGCTCAG TTGTGATTTA AAAGTGTTTT AAGTGATTCA	1080
TCCAAGGTTG TCTCCATGGA CTCCGGAAAT AACCACAGCC AAAATACCCC GGCGAGAGCA	1140
GAATATGAGA GGCAGATAGC ACCTATTATT ACTTTTATCT GGGTGCAGCC GGGCTCCAGA	1200
TTAATGGGGG CAAGCAAAGG TCCAGGCCTC AGACATGGCC AATTTTAAGC AACTAGTGTC	1260
TAGTGTGAT AAGTATTCAA TAAGGGCTTG GAATCTAATT ATTTGTAAGA AGTGTCTTTG	1320
TTTCCTGATC TGTAAGGAG ATGGGGAAAT ACTGTGATTT CGTGCCATTT TAAAAAAAT	1380
GAAAAAGCAA CATAAAAAAA TGCAAATACA GAAGAAAAGA TGTGGGGTTT GAAGTTGGCT	1440
TTAGAGAGCT GTGTAACCTT GGGGAAGTCA CCTCGTCTCT CTGTGACTTG AATCCCTAAT	1500
CAACAATAAT GCTGGGGGAC AAGGTGCTCT CCAAATCCCC CTTGAGATCA GAAAGTCTGG	1560
GATAACTATC CCAGATAGAT CCAAAGTTGT AAAGGAATCA AAAGGAATAG TCTAGTGTGA	1620
CAGATAACAC GGACTTGAAA CTTAGTTCAC CTGATATTTT TTAATGACTT AACACCACTA	1680

TACCTCACCC CAGAGCTAGC CTGTCCCCTG TGTGCCACC AGTAACCTTG AAATGTCGAC 1740

## (2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1385 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW093

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 8..1357

(ix) FEATURE:  
 (A) NAME/KEY: misc\_feature  
 (B) LOCATION: 798..1385  
 (D) OTHER INFORMATION: /label= SAC\_24538

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GTCTGACT	CTG	CCT	CAG	CCG	CTG	CCG	CCG	CCA	CTG	CTG	CTG	CAG	GAT	CCG	49	
	Leu	Pro	Gln	Pro	Leu	Pro	Pro	Pro	Leu	Leu	Leu	Gln	Asp	Pro		
	1				5					10						
CCG	CAG	CCA	CCA	GTC	TCG	CGC	GTC	GTC	GCT	CCC	TCC	TTG	GAC	AGT	AAT	97
Pro	Gln	Pro	Pro	Val	Ser	Arg	Val	Val	Ala	Pro	Ser	Leu	Asp	Ser	Asn	
15				20					25						30	
TTT	ATG	AAT	AAG	CAT	CAG	AAG	CCG	GTA	CTA	ACA	GGC	CAG	CGG	TTC	AAA	145
Phe	Met	Asn	Lys	His	Gln	Lys	Pro	Val	Leu	Thr	Gly	Gln	Arg	Phe	Lys	
			35					40						45		
ACC	CGG	AAA	AGG	GAT	GAA	AAA	GAG	AAA	TTC	GAA	CCC	ACA	GTC	TTC	AGG	193
Thr	Arg	Lys	Arg	Asp	Glu	Lys	Glu	Lys	Phe	Glu	Pro	Thr	Val	Phe	Arg	
			50				55						60			
GAT	ACA	CTT	GTC	CAG	GGG	CTT	AAT	GAA	GCT	GGT	GAT	GAC	CTT	GAA	GCT	241
Asp	Thr	Leu	Val	Gln	Gly	Leu	Asn	Glu	Ala	Gly	Asp	Asp	Leu	Glu	Ala	
	65					70						75				
GTA	GCC	AAA	TTC	TTG	GAC	TCT	ACT	GGC	TCA	CGA	CTA	GAT	TAT	CGT	CGC	289
Val	Ala	Lys	Phe	Leu	Asp	Ser	Thr	Gly	Ser	Arg	Leu	Asp	Tyr	Arg	Arg	
	80					85					90					
TAT	GCA	GAC	ACA	CTC	TTT	GAT	ATC	CTG	GTG	GCT	GGC	AGC	ATG	CTT	GCC	337
Tyr	Ala	Asp	Thr	Leu	Phe	Asp	Ile	Leu	Val	Ala	Gly	Ser	Met	Leu	Ala	
	95				100				105					110		
CCT	GGA	GGA	ACA	CGC	ATA	GAT	GAT	GGT	GAC	AAG	ACC	AAG	ATG	ACC	AAC	385
Pro	Gly	Gly	Thr	Arg	Ile	Asp	Asp	Gly	Asp	Lys	Thr	Lys	Met	Thr	Asn	
			115					120						125		

CAC TGT GTG TTT TCA GCA AAT GAA GAT CAT GAA ACC ATC CGA AAC TAT	433
His Cys Val Phe Ser Ala Asn Glu Asp His Glu Thr Ile Arg Asn Tyr	
130 135 140	
GCT CAG GTC TTC AAC AAA CTC ATC AGG AGA TAC AAA TAT TTG GAA AAG	481
Ala Gln Val Phe Asn Lys Leu Ile Arg Arg Tyr Lys Tyr Leu Glu Lys	
145 150 155	
GCA TTT GAA GAT GAA ATG AAA AAG CTT CTC CTC TTC CTT AAA GCA TTT	529
Ala Phe Glu Asp Glu Met Lys Lys Leu Leu Leu Phe Leu Lys Ala Phe	
160 165 170	
TCT GAA GCA GAG CAG ACA AAG TTG GCA ATG CTG TCT GGG ATC CTG TTG	577
Ser Glu Ala Glu Gln Thr Lys Leu Ala Met Leu Ser Gly Ile Leu Leu	
175 180 185 190	
GGC AAT GGA ACC CTC CCG GCC ACC ATC CTT ACC AGT CTC TTC ACC GAC	625
Gly Asn Gly Thr Leu Pro Ala Thr Ile Leu Thr Ser Leu Phe Thr Asp	
195 200 205	
AGC TTA GTC AAA GAA GGC ATC GCA GCC TCG TTC GCT GTC AAG CTT TTC	673
Ser Leu Val Lys Glu Gly Ile Ala Ala Ser Phe Ala Val Lys Leu Phe	
210 215 220	
AAA GCC TGG ATG GCA GAG AAA GAT GCC AAT TCT GTC ACC TCT TCA TTG	721
Lys Ala Trp Met Ala Glu Lys Asp Ala Asn Ser Val Thr Ser Ser Leu	
225 230 235	
AGA AAA GCC AAC CTG GAC AAG AGG CTG CTT GAA CTC TTC CCA GTG AAC	769
Arg Lys Ala Asn Leu Asp Lys Arg Leu Leu Glu Leu Phe Pro Val Asn	
240 245 250	
AGA CAG AGT GTG GAT CAT TTC GCC AAG TAC TTC ACT GAC GCA GGA CTG	817
Arg Gln Ser Val Asp His Phe Ala Lys Tyr Phe Thr Asp Ala Gly Leu	
255 260 265 270	
AAG GAG CTG TCA GAC TTT CTC CGA GTC CAG CAG TCA CTG GGC ACC AGG	865
Lys Glu Leu Ser Asp Phe Leu Arg Val Gln Gln Ser Leu Gly Thr Arg	
275 280 285	
AAG GAA CTG CAG AAG GAG CTG CAA GAG CGA CTG TCT CAG GAA TGC CCC	913
Lys Glu Leu Gln Lys Glu Leu Gln Glu Arg Leu Ser Gln Glu Cys Pro	
290 295 300	
ATC AAG GAG GTG GTG CTT TAC GTC AAA GAG GAA ATG AAA AGG AAC GAC	961
Ile Lys Glu Val Val Leu Tyr Val Lys Glu Glu Met Lys Arg Asn Asp	
305 310 315	
CTT CCG GAG ACA GCT GTG ATC GGG CTG CTG TGG ACC TGC GTC ATG AAC	1009
Leu Pro Glu Thr Ala Val Ile Gly Leu Leu Trp Thr Cys Val Met Asn	
320 325 330	
GCT GTG GAA TGG AAC AAG AAG GAG GAG CTG GTT GCT GAG CAG GCT CTT	1057
Ala Val Glu Trp Asn Lys Lys Glu Glu Leu Val Ala Glu Gln Ala Leu	
335 340 345 350	
AAG CAC CTG AAG CAA TAC GCT CCC CTG CTG GCC GTG TTC AGT TCC CAA	1105
Lys His Leu Lys Gln Tyr Ala Pro Leu Leu Ala Val Phe Ser Ser Gln	
355 360 365	
GGC CAG TCA GAG CTT GTC CTC CTG CAG AAG GTT CAG GAA TAC TGC TAT	1153
Gly Gln Ser Glu Leu Val Leu Leu Gln Lys Val Gln Glu Tyr Cys Tyr	

370	375	380	
GAC AAC ATC CAC TTC ATG AAA GCC TTC CAG AAG ATC GTG GTT CTC TTT Asp Asn Ile His Phe Met Lys Ala Phe Gln Lys Ile Val Val Leu Phe 385 390 395			1201
TAT AAA GCC GAC GTT CTG AGC GAG GAG GCG ATA CTG AAG TGG TAT AAA Tyr Lys Ala Asp Val Leu Ser Glu Glu Ala Ile Leu Lys Trp Tyr Lys 400 405 410			1249
GAA GCA CAT GCC GCC AAA GGC AAA AGC GTC TTC CTT GAC CAG ATG AAG Glu Ala His Ala Ala Lys Gly Lys Ser Val Phe Leu Asp Gln Met Lys 415 420 425 430			1297
AAA TTT GTG GAG TGG TTA CAA AAT GCA GAA GAA GAA TCT GAA TCA GAA Lys Phe Val Glu Trp Leu Gln Asn Ala Glu Glu Glu Ser Glu Ser Glu 435 440 445			1345
GGT GAA GAA AGC TAAATGTCTC CAGCACACTG CCTAAAAA Gly Glu Glu Ser 450			1385

## (2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 450 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Leu Pro Gln Pro Leu Pro Pro Pro Leu Leu Leu Gln Asp Pro Pro Gln 1 5 10 15
Pro Pro Val Ser Arg Val Val Ala Pro Ser Leu Asp Ser Asn Phe Met 20 25 30
Asn Lys His Gln Lys Pro Val Leu Thr Gly Gln Arg Phe Lys Thr Arg 35 40 45
Lys Arg Asp Glu Lys Glu Lys Phe Glu Pro Thr Val Phe Arg Asp Thr 50 55 60
Leu Val Gln Gly Leu Asn Glu Ala Gly Asp Asp Leu Glu Ala Val Ala 65 70 75 80
Lys Phe Leu Asp Ser Thr Gly Ser Arg Leu Asp Tyr Arg Arg Tyr Ala 85 90 95
Asp Thr Leu Phe Asp Ile Leu Val Ala Gly Ser Met Leu Ala Pro Gly 100 105 110
Gly Thr Arg Ile Asp Asp Gly Asp Lys Thr Lys Met Thr Asn His Cys 115 120 125
Val Phe Ser Ala Asn Glu Asp His Glu Thr Ile Arg Asn Tyr Ala Gln 130 135 140
Val Phe Asn Lys Leu Ile Arg Arg Tyr Lys Tyr Leu Glu Lys Ala Phe 145 150 155 160



Glu Asp Glu Met Lys Lys Leu Leu Leu Phe Leu Lys Ala Phe Ser Glu  
 165 170 175  
 Ala Glu Gln Thr Lys Leu Ala Met Leu Ser Gly Ile Leu Leu Gly Asn  
 180 185 190  
 Gly Thr Leu Pro Ala Thr Ile Leu Thr Ser Leu Phe Thr Asp Ser Leu  
 195 200 205  
 Val Lys Glu Gly Ile Ala Ala Ser Phe Ala Val Lys Leu Phe Lys Ala  
 210 215 220  
 Trp Met Ala Glu Lys Asp Ala Asn Ser Val Thr Ser Ser Leu Arg Lys  
 225 230 235 240  
 Ala Asn Leu Asp Lys Arg Leu Leu Glu Leu Phe Pro Val Asn Arg Gln  
 245 250 255  
 Ser Val Asp His Phe Ala Lys Tyr Phe Thr Asp Ala Gly Leu Lys Glu  
 260 265 270  
 Leu Ser Asp Phe Leu Arg Val Gln Gln Ser Leu Gly Thr Arg Lys Glu  
 275 280 285  
 Leu Gln Lys Glu Leu Gln Glu Arg Leu Ser Gln Glu Cys Pro Ile Lys  
 290 295 300  
 Glu Val Val Leu Tyr Val Lys Glu Glu Met Lys Arg Asn Asp Leu Pro  
 305 310 315 320  
 Glu Thr Ala Val Ile Gly Leu Leu Trp Thr Cys Val Met Asn Ala Val  
 325 330 335  
 Glu Trp Asn Lys Lys Glu Glu Leu Val Ala Glu Gln Ala Leu Lys His  
 340 345 350  
 Leu Lys Gln Tyr Ala Pro Leu Leu Ala Val Phe Ser Ser Gln Gly Gln  
 355 360 365  
 Ser Glu Leu Val Leu Leu Gln Lys Val Gln Glu Tyr Cys Tyr Asp Asn  
 370 375 380  
 Ile His Phe Met Lys Ala Phe Gln Lys Ile Val Val Leu Phe Tyr Lys  
 385 390 395 400  
 Ala Asp Val Leu Ser Glu Glu Ala Ile Leu Lys Trp Tyr Lys Glu Ala  
 405 410 415  
 His Ala Ala Lys Gly Lys Ser Val Phe Leu Asp Gln Met Lys Lys Phe  
 420 425 430  
 Val Glu Trp Leu Gln Asn Ala Glu Glu Glu Ser Glu Ser Glu Gly Glu  
 435 440 445  
 Glu Ser  
 450

## (2) INFORMATION FOR SEQ ID NO:85:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1463 base pairs  
 (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW094

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 255..1238

(ix) FEATURE:

(A) NAME/KEY: misc\_feature

(B) LOCATION: 1050..1355

(D) OTHER INFORMATION: /label= SAC\_23896

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

GTCGACCCCG AACACCCACA CTCGGTCCCC AGCCGTTTTT CTCCCAGAGA TCCGGAAGCC	60
CCGACGTCGT GCCCTACACC AAAGATCGAG AAAACCACTC AGACGCGGCT CTTAGGCGGC	120
AGTTCGGTTC CCATAGCAAC CCTCGGAGCA CCCGCATCCC GCTTCCGGCC GGGTGGCTCG	180
GGCGCTCCCC TGACGTCACG GCGCCGCTTT CCGGTGACTG AGTCGAGCTG AGGTTGTTGT	240
GGGCCGGGGG CGCC ATG GGG GCG ACT GGC GAC ACC GAG CAG CCG CGG GGC	290
Met Gly Ala Thr Gly Asp Thr Glu Gln Pro Arg Gly	
1 5 10	
CCC GGC GGG GCG GAG CGA GGC GGC CTG GAG CTG GGC GAC GCG GGC GCG	338
Pro Gly Gly Ala Glu Arg Gly Gly Leu Glu Leu Gly Asp Ala Gly Ala	
15 20 25	
GCG GGC CAG CCG GTT CTC ACG AAC CCT TGG AAC ATA ATG ATC AAA CAC	386
Ala Gly Gln Pro Val Leu Thr Asn Pro Trp Asn Ile Met Ile Lys His	
30 35 40	
CGG CAG GTG CAG CGA AGA GGC CGC CGA TCT CAG ATG ACC ACA AGT TTC	434
Arg Gln Val Gln Arg Arg Gly Arg Arg Ser Gln Met Thr Thr Ser Phe	
45 50 55 60	
ACA GAT CCA GCC ATC TCT ATG GAT CTC CTC CGT GCT GTC CTG CAG CCT	482
Thr Asp Pro Ala Ile Ser Met Asp Leu Leu Arg Ala Val Leu Gln Pro	
65 70 75	
AGC ATC AAT GAG GAG ATC CAG AGT GTC TTC AAC AAG TAC ATG AAG TTC	530
Ser Ile Asn Glu Ile Gln Ser Val Phe Asn Lys Tyr Met Lys Phe	
80 85 90	
TTC CAG AAG GCA GCG CTG AAT GTG CGA GAC AAC GTT GGT GAA GAG GTG	578
Phe Gln Lys Ala Ala Leu Asn Val Arg Asp Asn Val Gly Glu Glu Val	
95 100 105	
GAT GCA GAA CAG TTG ATT CAG GAG GCC TGC CGC AGC TGC CTG GAG CAG	626
Asp Ala Glu Gln Leu Ile Gln Glu Ala Cys Arg Ser Cys Leu Glu Gln	
110 115 120	

GCA AAG CTA CTC TTT TCA GAT GGA GAA AAA GTG ATA CCC AGA TTG GCC Ala Lys Leu Leu Phe Ser Asp Gly Glu Lys Val Ile Pro Arg Leu Ala 125 130 135 140	674
CAT GAA CTT CCA GGG ATC AAG CGT GGC CGG CAG GCA GAA GAG GAG TCC His Glu Leu Pro Gly Ile Lys Arg Gly Arg Gln Ala Glu Glu Glu Ser 145 150 155	722
CAC CGA GGA AGC CCC ATT CCC AAA AAG AGG AAA GGT CGG CCT CCT GGA His Arg Gly Ser Pro Ile Pro Lys Lys Arg Lys Gly Arg Pro Pro Gly 160 165 170	770
CAT GTC CTG TCA AAT GAC CGC GCA GCT GCT GGC ATG GTA TGG AAA CCA His Val Leu Ser Asn Asp Arg Ala Ala Ala Gly Met Val Trp Lys Pro 175 180 185	818
AAA TCC TGT GAA CCA ATT CGC CGA GAA GGC CCC AAG TGG GAC CCA GCT Lys Ser Cys Glu Pro Ile Arg Arg Glu Gly Pro Lys Trp Asp Pro Ala 190 195 200	866
CGG CTG AAT GAA TCT ACC ACC TTT GTT TTG GGG TCT CGA GCC AAC AAG Arg Leu Asn Glu Ser Thr Thr Phe Val Leu Gly Ser Arg Ala Asn Lys 205 210 215 220	914
GCC TTA GGG ATG GGA GGC ACC AGA GGG AGG ATT TAC ATC AAG CAC CCA Ala Leu Gly Met Gly Gly Thr Arg Gly Arg Ile Tyr Ile Lys His Pro 225 230 235	962
CAC CTC TTT AAG TAT GCA GCA GAT CCT CAG GAC AAG CAC TGG CTG GCT His Leu Phe Lys Tyr Ala Ala Asp Pro Gln Asp Lys His Trp Leu Ala 240 245 250	1010
GAG CAG CAT CAT ATG CGG GCA ACA GGA GGA AAG ATG GCG TAC CTT CTC Glu Gln His His Met Arg Ala Thr Gly Gly Lys Met Ala Tyr Leu Leu 255 260 265	1058
ATT GAG GAA GAC ATC CGA GAC TTG GCT GCC AGC GAT GAC TAC AGA GGA Ile Glu Glu Asp Ile Arg Asp Leu Ala Ala Ser Asp Asp Tyr Arg Gly 270 275 280	1106
TGC TTG GAC CTG AAG TTG GAG GAG CTG AAG TCC TTT GTT TTG CCA TCC Cys Leu Asp Leu Lys Leu Glu Glu Leu Lys Ser Phe Val Leu Pro Ser 285 290 295 300	1154
TGG ATG GTT GAG AAG ATG CGG AAA TAC ATG GAG ACA CTA CGG ACA GAA Trp Met Val Glu Lys Met Arg Lys Tyr Met Glu Thr Leu Arg Thr Glu 305 310 315	1202
AAT GAG CAT CGT GCT GCT GAA GCA ACT CCC CAG ACC TGAGCCGAGT Asn Glu His Arg Ala Ala Glu Ala Thr Pro Gln Thr 320 325	1248
GTCTGGCTA CTACACTTGG CAGTCTGCCT CCCAGACCCT CTTTCCCCGC CCGGCTGAGG	1308
CCATCATGGG GATGCGGTCT AGTTGGCTCT TAGCAGCATC AAGCTGTACA TGAGCTAGTT	1368
TGTAGTGACT CACTGCAGAG CCCCCAGAC TGGCTTGTGG TTCTGTTTCT AAAGTTATTG	1428
GAATAAGAAG CAATTAAACA AGTTTGTAAAT AAAAA	1463

(2) INFORMATION FOR SEQ ID NO:86:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 328 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

```

Met Gly Ala Thr Gly Asp Thr Glu Gln Pro Arg Gly Pro Gly Gly Ala
 1           5           10           15
Glu Arg Gly Gly Leu Glu Leu Gly Asp Ala Gly Ala Ala Gly Gln Pro
          20           25           30
Val Leu Thr Asn Pro Trp Asn Ile Met Ile Lys His Arg Gln Val Gln
          35           40           45
Arg Arg Gly Arg Arg Ser Gln Met Thr Thr Ser Phe Thr Asp Pro Ala
          50           55           60
Ile Ser Met Asp Leu Leu Arg Ala Val Leu Gln Pro Ser Ile Asn Glu
          65           70           75           80
Glu Ile Gln Ser Val Phe Asn Lys Tyr Met Lys Phe Phe Gln Lys Ala
          85           90           95
Ala Leu Asn Val Arg Asp Asn Val Gly Glu Glu Val Asp Ala Glu Gln
          100          105          110
Leu Ile Gln Glu Ala Cys Arg Ser Cys Leu Glu Gln Ala Lys Leu Leu
          115          120          125
Phe Ser Asp Gly Glu Lys Val Ile Pro Arg Leu Ala His Glu Leu Pro
          130          135          140
Gly Ile Lys Arg Gly Arg Gln Ala Glu Glu Glu Ser His Arg Gly Ser
          145          150          155          160
Pro Ile Pro Lys Lys Arg Lys Gly Arg Pro Pro Gly His Val Leu Ser
          165          170          175
Asn Asp Arg Ala Ala Ala Gly Met Val Trp Lys Pro Lys Ser Cys Glu
          180          185          190
Pro Ile Arg Arg Glu Gly Pro Lys Trp Asp Pro Ala Arg Leu Asn Glu
          195          200          205
Ser Thr Thr Phe Val Leu Gly Ser Arg Ala Asn Lys Ala Leu Gly Met
          210          215          220
Gly Gly Thr Arg Gly Arg Ile Tyr Ile Lys His Pro His Leu Phe Lys
          225          230          235          240
Tyr Ala Ala Asp Pro Gln Asp Lys His Trp Leu Ala Glu Gln His His
          245          250          255
Met Arg Ala Thr Gly Gly Lys Met Ala Tyr Leu Leu Ile Glu Glu Asp
          260          265          270
Ile Arg Asp Leu Ala Ala Ser Asp Asp Tyr Arg Gly Cys Leu Asp Leu
          275          280          285

```

Lys Leu Glu Glu Leu Lys Ser Phe Val Leu Pro Ser Trp Met Val Glu  
290 295 300

Lys Met Arg Lys Tyr Met Glu Thr Leu Arg Thr Glu Asn Glu His Arg  
305 310 315 320

Ala Ala Glu Ala Thr Pro Gln Thr  
325

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1276 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: HW095

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 7..876

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: complement (631..798)
- (D) OTHER INFORMATION: /label= SAC\_23802

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

GTCGAC ATT TTC TCA GAA TCC TAC GGA GGA AAG ATG GCT GCT GGC ATC Ile Phe Ser Glu Ser Tyr Gly Gly Lys Met Ala Ala Gly Ile 1 5 10	48
AGT TTA GAA CTT CAC AAG GCT ATT CAG CAA GGG ACC ATC AAG TGC AAC Ser Leu Glu Leu His Lys Ala Ile Gln Gln Gly Thr Ile Lys Cys Asn 15 20 25 30	96
TTC TCT GGG GTT GCT TTG GGT GAC TCC TGG ATC TCC CCT GTG GAT TCA Phe Ser Gly Val Ala Leu Gly Asp Ser Trp Ile Ser Pro Val Asp Ser 35 40 45	144
GTG CTG TCC TGG GGA CCT TAC CTG TAC AGC GTG TCT CTC CTT GAT AAT Val Leu Ser Trp Gly Pro Tyr Leu Tyr Ser Val Ser Leu Leu Asp Asn 50 55 60	192
AAA GGC TTG GCT GAG GTG TCC GAC ATT GCG GAG CAA GTC CTC AAT GCT Lys Gly Leu Ala Glu Val Ser Asp Ile Ala Glu Gln Val Leu Asn Ala 65 70 75	240
GTA AAC AAG GGC TTC TAC AAG GAA GCC ACT CAG CTG TGG GGG AAA GCA Val Asn Lys Gly Phe Tyr Lys Glu Ala Thr Gln Leu Trp Gly Lys Ala 80 85 90	288
GAA ATG ATC ATT GAA AAG AAC ACC GAC GGG GTA AAC TTC TAT AAC ATC Glu Met Ile Ile Glu Lys Asn Thr Asp Gly Val Asn Phe Tyr Asn Ile	336

95	100	105	110	
TTA ACT AAA AGC ACC CCC GAC ACC TCT ATG GAG TCG AGC CTC GAG TTC Leu Thr Lys Ser Pro Asp Thr Ser Met Glu Ser Ser Leu Glu Phe 115 120 125				384
TTC CGG AGC CCC TTA GTT CGT CTC TGT CAG CGC CAC GTG AGA CAC CTA Phe Arg Ser Pro Leu Val Arg Leu Cys Gln Arg His Val Arg His Leu 130 135 140				432
CAA GGA GAC GCC TTA AGT CAG CTC ATG AAC GGT CCC ATC AAA AAG AAG Gln Gly Asp Ala Leu Ser Gln Leu Met Asn Gly Pro Ile Lys Lys Lys 145 150 155				480
CTC AAA ATT ATC CCT GAC GAT GTC TCC TGG GGA GCC CAG TCA TCC TCC Leu Lys Ile Ile Pro Asp Asp Val Ser Trp Gly Ala Gln Ser Ser Ser 160 165 170				528
GTC TTC ATA AGC ATG GAA GAG GAC TTC ATG AAG CCT GTC ATC GAC ATC Val Phe Ile Ser Met Glu Glu Asp Phe Met Lys Pro Val Ile Asp Ile 175 180 185 190				576
GTG GAT ACG TTG CTG GAA CTC GGG GTC AAT GTG ACT GTG TAC AAT GGG Val Asp Thr Leu Leu Glu Leu Gly Val Asn Val Thr Val Tyr Asn Gly 195 200 205				624
CAG CTG GAT CTC ATT GTG GAC ACC ATA GGT CAG GAG TCC TGG GTT CAG Gln Leu Asp Leu Ile Val Asp Thr Ile Gly Gln Glu Ser Trp Val Gln 210 215 220				672
AAG CTG AAG TGG CCA CAG CTG TCC AGA TTC AAT CAG CTA AAA TGG AAG Lys Leu Lys Trp Pro Gln Leu Ser Arg Phe Asn Gln Leu Lys Trp Lys 225 230 235				720
GCC CTG TAC ACC AAT CCT AAG TCT TCA GAA ACA TCT GCG TTT GTC AAG Ala Leu Tyr Thr Asn Pro Lys Ser Ser Glu Thr Ser Ala Phe Val Lys 240 245 250				768
TCC TAT GAG AAC CTA GCG TTC TAC TGG ATC CTA AAG GCG GGT CAC ATG Ser Tyr Glu Asn Leu Ala Phe Tyr Trp Ile Leu Lys Ala Gly His Met 255 260 265 270				816
GTT CCT GCT GAC CAA GGG GAC ATG GCT CTG AAG ATG ATG AGG CTG GTT Val Pro Ala Asp Gln Gly Asp Met Ala Leu Lys Met Met Arg Leu Val 275 280 285				864
ACT CAG CAG GAG TAGCTGAGCT GAGCTGGCCC TGGAGGCCCT GGAGGCCCTG Thr Gln Gln Glu 290				916
GAGGCCCTGG AGTAGGGCCC AGGATGCAGG TGCTAATGTC TATCCCCGGC GCTCTTCTTC				976
CCGACTCTAC CATGGGATGT AACTCCAGGA GCCCTTGCCA TCTCCGGTAC CAAAAGACTG				1036
TGGCTTCCGT GTCTACTCAG AAATCAGTTC TACTTCGTAA ACAGTGTTTA AAACCAGACT				1096
CATTTAATCA GAGTGAAGGA TTGCAGTCCA TTGGCTTCTT AGCACAGAAG CAGCTGATAA				1156
CACAAGTAAA CCCCAGCCCT TGAGAGGTAG AAGCAAGAGG ATCAGAGGTT CAAGCGCATC				1216
CTCGGCTCCA TCACAAGTTC AAAAGCCGCC TGCACCAAAT GGGAGTCCTT GTCTCAAAAA				1276

WO 98/53071

## (2) INFORMATION FOR SEQ ID NO:88:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 290 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

```

Ile Phe Ser Glu Ser Tyr Gly Gly Lys Met Ala Ala Gly Ile Ser Leu
 1           5           10           15
Glu Leu His Lys Ala Ile Gln Gln Gly Thr Ile Lys Cys Asn Phe Ser
 20           25           30
Gly Val Ala Leu Gly Asp Ser Trp Ile Ser Pro Val Asp Ser Val Leu
 35           40           45
Ser Trp Gly Pro Tyr Leu Tyr Ser Val Ser Leu Leu Asp Asn Lys Gly
 50           55           60
Leu Ala Glu Val Ser Asp Ile Ala Glu Gln Val Leu Asn Ala Val Asn
 65           70           75           80
Lys Gly Phe Tyr Lys Glu Ala Thr Gln Leu Trp Gly Lys Ala Glu Met
 85           90           95
Ile Ile Glu Lys Asn Thr Asp Gly Val Asn Phe Tyr Asn Ile Leu Thr
100           105           110
Lys Ser Thr Pro Asp Thr Ser Met Glu Ser Ser Leu Glu Phe Phe Arg
115           120           125
Ser Pro Leu Val Arg Leu Cys Gln Arg His Val Arg His Leu Gln Gly
130           135           140
Asp Ala Leu Ser Gln Leu Met Asn Gly Pro Ile Lys Lys Lys Leu Lys
145           150           155           160
Ile Ile Pro Asp Asp Val Ser Trp Gly Ala Gln Ser Ser Ser Val Phe
165           170           175
Ile Ser Met Glu Glu Asp Phe Met Lys Pro Val Ile Asp Ile Val Asp
180           185           190
Thr Leu Leu Glu Leu Gly Val Asn Val Thr Val Tyr Asn Gly Gln Leu
195           200           205
Asp Leu Ile Val Asp Thr Ile Gly Gln Glu Ser Trp Val Gln Lys Leu
210           215           220
Lys Trp Pro Gln Leu Ser Arg Phe Asn Gln Leu Lys Trp Lys Ala Leu
225           230           235           240
Tyr Thr Asn Pro Lys Ser Ser Glu Thr Ser Ala Phe Val Lys Ser Tyr
245           250           255
Glu Asn Leu Ala Phe Tyr Trp Ile Leu Lys Ala Gly His Met Val Pro
260           265           270

```

Ala Asp Gln Gly Asp Met Ala Leu Lys Met Met Arg Leu Val Thr Gln  
 275 280 285

Gln Glu  
 290

(2) INFORMATION FOR SEQ ID NO:89:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 2421 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: HW096

(ix) FEATURE:  
 (A) NAME/KEY: CDS  
 (B) LOCATION: 295..1302

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89;

GTCGACGTCT GATTTTACAT TCATGCAGTT GGGAGTAAGA GGAAAAAGAT TCTGAAGTGT	60
GGGGAACCTC ACAAAGAGGG GAACAAAGTC TGTGTCTATG GCTCCAAGGG AGAGAACATC	120
AGAGACACTC TGAGGAAGGA CGGCAGGTTT TGTCTTTTCA TAGAGAATAA CCATTGGAAA	180
CTCATTAATG ACCTGGACAC TATCATAGAA AACACTCAGC CATTTGATGA GTTGGAGGGC	240
AAGCTCTTTC AGGTTGCAGC TGAGCTACCA AAGAACCCTA GGGTAGTCTC TGTC ATG	297
	Met 1
CAG AAT TCT GGG TTA GAG AAC AGA GAC TTC CAT AAG CTA GAA GAC TAT	345
Gln Asn Ser Gly Leu Glu Asn Arg Asp Phe His Lys Leu Glu Asp Tyr	5 10 15
ATT GTG AAT GCG AAC CCT ACA TTG AAA GAA GAA GGA GAA AAA CTC AGA	393
Ile Val Asn Ala Asn Pro Thr Leu Lys Glu Glu Gly Glu Lys Leu Arg	20 25 30
GCA TAC ATC AAG AAA GAA AGT AAA AAA AGA AAG AAA GGT TCC TTA TTC	441
Ala Tyr Ile Lys Lys Glu Ser Lys Lys Arg Lys Lys Gly Ser Leu Phe	35 40 45
AAA GTG CAT AAA GAA CAC TTT GGG AAA ATA ACA AAA AAT TCT ACT CCT	489
Lys Val His Lys Glu His Phe Gly Lys Ile Thr Lys Asn Ser Thr Pro	50 55 60 65
GTT AAA GTG CAC AAA CAT CTT TCG AAA GTC AGT GAT TCA GTT GGG TTC	537
Val Lys Val His Lys His Leu Ser Lys Val Ser Asp Ser Val Gly Phe	70 75 80
CTA TGG TGG AAC AAC AAT GGA AAG GAG GGC TGT GCC ACC TGC TTT GTT	585
Leu Trp Trp Asn Asn Asn Gly Lys Glu Gly Cys Ala Thr Cys Phe Val	



85	90	95	
TTT AAA GGA TTA TAC ATT TTG ACT TGC CGG CAT GTG ATA ACT AGT ATT Phe Lys Gly Leu Tyr Ile Leu Thr Cys Arg His Val Ile Thr Ser Ile 100 105 110			633
GTG GGT GAA GGC ATA GAT CCA TGT CAG TGG GCA AGC TTA ATT AGT CAG Val Gly Glu Gly Ile Asp Pro Cys Gln Trp Ala Ser Leu Ile Ser Gln 115 120 125			681
TGT GTA AAG GTG ACC TTT GGT TAT GAA GAG TTC CCA CTA GTA GAA GAC Cys Val Lys Val Thr Phe Gly Tyr Glu Glu Phe Pro Leu Val Glu Asp 130 135 140 145			729
AAG TTT TTT GAG GTT AAA CCT TGG TTT GAG ATA TCT AAT GAA AGC CTT Lys Phe Phe Glu Val Lys Pro Trp Phe Glu Ile Ser Asn Glu Ser Leu 150 155 160			777
GAC TAT GCT GTC CTA GAA CTG AGG GAA AAT GGA CAA GAA GTT CCT GCT Asp Tyr Ala Val Leu Glu Leu Arg Glu Asn Gly Gln Glu Val Pro Ala 165 170 175			825
GGA CTG TAT AAT GGA ATA GGA CCT GTG CCA CTT GGT GGG TTA ATA TAT Gly Leu Tyr Asn Gly Ile Gly Pro Val Pro Leu Gly Gly Leu Ile Tyr 180 185 190			873
ATC ATT GGC CAT CCT GAT GGA GAA AAG AAG TCT ACT GAT GGC TGT GCA Ile Ile Gly His Pro Asp Gly Glu Lys Lys Ser Thr Asp Gly Cys Ala 195 200 205			921
GTG GTC CCT CAA AGT CAT AGA GGG CAA AAA TGT CAG GAA AAT TTT CAA Val Val Pro Gln Ser His Arg Gly Gln Lys Cys Gln Glu Asn Phe Gln 210 215 220 225			969
GCA AGA GAG AAA GCA AAC TAC TGG ATT TCT ACA TCT TTT TTT CCT ATA Ala Arg Glu Lys Ala Asn Tyr Trp Ile Ser Thr Ser Phe Phe Pro Ile 230 235 240			1017
TAC ACA CAA AGA AGT TTC CAC GAA GTC TTT CAC AAC CAT GAT GTG GTT Tyr Thr Gln Arg Ser Phe His Glu Val Phe His Asn His Asp Val Val 245 250 255			1065
ACT TAT GAC ACC ACT TTT TTT GGT GGG TCT TCT GGA TCC CCA GTA TTT Thr Tyr Asp Thr Thr Phe Phe Gly Gly Ser Ser Gly Ser Pro Val Phe 260 265 270			1113
GAT TCT AAT GGT TCA TTG GTG GCC ATG CAT GCT GCT GGC ATC ACT TGT Asp Ser Asn Gly Ser Leu Val Ala Met His Ala Ala Gly Ile Thr Cys 275 280 285			1161
ACA GAC CAG AAT GGA GTT TTT AAT ATC ATT GAG TTT GGT TTT ACT ATG Thr Asp Gln Asn Gly Val Phe Asn Ile Ile Glu Phe Gly Phe Thr Met 290 295 300 305			1209
GAA TCC ATT CTT GCT AAT ATT AAG CAA ACT AAA GAG TGG TAT GAT ACC Glu Ser Ile Leu Ala Asn Ile Lys Gln Thr Lys Glu Trp Tyr Asp Thr 310 315 320			1257
ATT TTT GTA AAT GAT CGG GAT GCA GAA ATG CTG AGC ACA GAC TCC Ile Phe Val Asn Asp Arg Asp Ala Glu Met Leu Ser Thr Asp Ser 325 330 335			1302
TGAGGACTGG AAAGGGTGTA TTCTAGTTCA CAACTTTAAG GGAATTGCCT AAGGTGTTTT			1362

TCTTCAGATA ATAATGTTTT ATGAAC TTTC AAAATGATTA ATTTCAACCA ATAATCATAG 1422  
 TTCTAGAGTG TATCTTAACT ATCAAGCATT TTCATGGTTT ATTTAAACAA ACAATCAAAA 1482  
 ACAAACTAT CAGTTGAACT ATTTGTGTAA AGTAGCATGA AAAAGAGGAT TAAGTGCTGA 1542  
 TGTGGGATCC GAGGGTCAGG AGATGCTAAG ATGTGAAAAA GTTGAGAGCA TAACTGGAG 1602  
 TTATTAGCCA GAAGATTCAT GGAAAAGAAA GAAGCCTACC ACGTTCCTTG AACTTTAAAC 1662  
 CAATTTCAAA CTCACTCATT TGGTTCTACC CATATCTTCC TTCTCTGCTG ACCAGAGCTC 1722  
 AGCCAGGACA CTGAATCACA GCGACTTAAC CCTCCAGGG CCCTGAACAT GGTGGCATT 1782  
 CTTCCATTTT GTCTTTTAA CTTAAGACTT AGATCTGAAA TGAACCCACA TGAAGTGGCA 1842  
 TGGAATCAT TGAGCAAGCA AGGAGCTGGT CTCTGCCACA GAATGCATGT ACCTGAGACT 1902  
 TTCTGGACTT GAGTGTGGA GAGCAAACAT TGGAATAGGT GGTTCCTACT TGATACGATT 1962  
 CACAATGGAG AGTGCTAGGA GGATATTGGG CACAACCTAG GACAATTCTC AGCAGTGTGA 2022  
 CACTCAAGTC TTGGTCCTTT GAAGAGAAAG AAACTGCGT TCAAGAAGTA ATCTCATCAA 2082  
 AAACAGAGTC AAGGAGAATT TATGAACAGA AGACCAGAGA TCTTAGATTA TTTGAGAGAA 2142  
 AGTAGGAATA GATGAGATGC TGAAGGAATG TTGCTGGGAT CCAGCGATTG CTTCAATTCT 2202  
 TCCTTTATCA TAACTGAAAT ATGTCCTGTA TGAGAGCTAA CATGATCCTA ATCCACAAC 2262  
 CCCCTTATTT TCCAGACATT TCCTAGTGCT ATTGTGTGAT ATTTGTGTGTC CCCTGCCCAA 2322  
 TGTATTTTTC CCAATCAAA TACAGGTGAA TCAATTAAAA AACAGTGTA TATCTGTAGT 2382  
 AGGTTTTGAA ATATAAGATG AAAAAATTAA AAAGTCGAC 2421

## (2) INFORMATION FOR SEQ ID NO:90:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 336 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Met Gln Asn Ser Gly Leu Glu Asn Arg Asp Phe His Lys Leu Glu Asp  
 1 5 10 15  
 Tyr Ile Val Asn Ala Asn Pro Thr Leu Lys Glu Glu Gly Glu Lys Leu  
 20 25 30  
 Arg Ala Tyr Ile Lys Lys Glu Ser Lys Lys Arg Lys Lys Gly Ser Leu  
 35 40 45  
 Phe Lys Val His Lys Glu His Phe Gly Lys Ile Thr Lys Asn Ser Thr  
 50 55 60  
 Pro Val Lys Val His Lys His Leu Ser Lys Val Ser Asp Ser Val Gly  
 65 70 75 80

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 836 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: HW097

(ix) FEATURE:  
(A) NAME/KEY: misc\_feature  
(B) LOCATION: complement (203..553)  
(D) OTHER INFORMATION: /label= SAC\_23978

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

GTCGACTATT ACTATAGCCT TGTATATACT CCCCTCCGGC CCTGTTTGTC TCAGTCCCAT	60
TCCCTCTTGT CTCTGAGAAT CATCCCCCAG CCCCAAGTTC TCCCTTCTGT CTCCCCCAC	120
CCCCCTGCC TCCCATTCTC CTGTCTCTGA ATGCCTTTGC CTGTATAAAG AGTTGGATTG	180
TCCCCTGGTG TCTGTACTGT GTACACACAT CCCTTTGAGA AGCACAAGGA GATGACACGC	240
GCATTGTAAC CTTCACTG TCTCGGTGGC GACATAAAGG AAGCTGTGAA TTACAAGCTC	300
TGCCTCTTTC TGGCCTCTCC CTCGGCCCCC ATGACCGGG CACTCTCTGC CCTCCCCACG	360
GCCTTAACAT GTTCCTGCTC CACCTATCCT GGTGCCCTTT GTCTGGCTGA CTGTAGCCTC	420
CCTGAGGGAA GGGATTGCTG GAGAGGTAGC TCCTATCCAG GAATGAAGGA CACTGGCTGG	480
AAACTAAGCC AAGAATGTTA CAGACAGGGG GAGCTGGGA TTGGTGACTG ACCCTTGACG	540
TTGGGCAGTG AGTACCCAAG GCCAGAAGCT GAAGCCCTCA TTCATTGAGT GGCCTTTCTT	600
TCCTCAGGGT GCCCAGTGCC CTGTCTCAC TTTGATGCCT AGGCCCACTG CTTGCATTTT	660
CTGGTGAGCC AGGGCTTGGT CATGTGAGGA AAGAGGTAGG CCTAGAGCAG GAGTTGAAAG	720
CCCTGCTGTT GTGTGTCCCT ACCATACAGG GCTGGCTAAA AGTCGAGGTA GACGCTTATC	780
TAGGAAGCGC AAGGCCCTGG GTTCGGTCCC CAGCTCGGAA AAAAAAGAAC CAAAAA	836

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 318 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC23798

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

GGGGGATCCG CCCGATCCAA GACTGCCCAG GGTGGCTGCT CCATCTCACC CAGAGCAACC	60
---	----

GGAGCTGGAC CTGGGGGTCG GAGCACTTCG AAGGTGATGG GTTATGGTGC CGCCTGCTGG 120  
ACATAGTTGT CCAGAGCTCT TCCTGAAAGG TTTGCTACAG GATTGCGCCT TCTCATCCAG 180  
TCTAAGAAAC TACTTAGTCA ACTCCTAGTG TCTAAGCAGA AGGCTTCAGC TATGGCTTCT 240  
CCCCTTGGGG AGGAAGAGGA AGAGGTACAA ATTCTTGTTT TGCTGCTCCC AGGGAAAAGA 300  
TCGGGCTAGA GCGGCCGC 318

## (2) INFORMATION FOR SEQ ID NO:93:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 278 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_23800

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

GGGGGATCCG CCCGATCACA CGGCTGCCTT CTGGTATACA GCAGAGGGGA ATGCCAAGGT 60  
ATAAGTATCC TTCAATTCCT TCTTTGCCAC ATACCAGCAT TTCTTTGCCC TTTTGGCATT 120  
CTACCTCTTC TTTGACAGCA CTGAGACTGT TAAGTGCTGC ATACAAGCTA CTTGTCTCCC 180  
TATCTGTNNG CACATGTTTC ATGTCATGAA GATGATAAGG TCAGGCAAAG AGGCATGAAG 240  
ATTCCTGAAT AGTGCCCGA TCGGGCTAGA GCGGCCGC 278

## (2) INFORMATION FOR SEQ ID NO:94:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 393 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_23801

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

GGGGGATCCG CCCGATCCAG TAGCGACAGC CCTGGCCCCCT GACGAGGATG GAAGTGACGT 60  
TGTAGCNATT GTCTTCCTGT AGCTCATAGA TGGTGCTGTN NATGGTAAAG CGGCTTTGTC 120

```

TTTCTTTCTG GACCGCATTG CCTGCCAGGN NGACAACGAA CCACCTGCCC TGGAACCGTT      180
CGGTCCAGAA GCCTGGCTGC AGGGGCACAC TGATCGATCC AGTAGCGACA GCCCTGGCCC      240
CTGACGAGGA TGGAAGTGAC GTTGTAGCTA TTGTCTTCCT GTAGCTCATA GATGGTGCTG      300
TACATGGTAA AGCGGCTTTG TCTTTCTTTC TGGACCGCAT TGCCTGCCAG GCCGACAACG      360
AACCACCTGC CCTGGAACCG TTCGGTCCAG AAG                                     393

```

## (2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 531 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
 (B) CLONE: SAC\_23833

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

```

GGGGGATCCG CCCGATCTGT ATCTCGTCCC TTGGCTGGCT GAGCCTCTGG AGGAGGTGTC      60
CAAAGCCCCG GTGGATTCTT GTCTCAGGGG TCTCTGTGAG ATTGAACTTG AGACCTTCTA      120
GAATCTCTTC CATGCTGTTG CCCTTTGCTC CCAGGGACAC GACGGCCAAG GCGGCTGAGA      180
TGCTAAGTGG GGAGAAGACA ACATTTTAT CTGGATTCCT CAAAGCCAGC TTCCTGTAGA      240
GGCTGAAGGC AAAGTCAGTA TTGATGGAAG CCAGTGTGAG ACTGTCCAGT TGTGTCCCCT      300
TGTCCTGGTC TTCATGGAAT AGAGTGTCCT CTCCCAGTAT GCAAATCTGG GAAGCAGAGG      360
AGAGCAGAGC CCGATCGATC AAGATGACTA AGATGCTCAA AGGATTCGAC GCTGTGGGAA      420
ATGCCACAGG TTTCCGGTTC GCCTACACCC CAGCTATGGA GAGCCTCTGT GGATATGTCA      480
CAAGTCCAGA ACCGCAGCGA GGAGTTTCTC ATCGCGGGCC GTTTAAGGAC G              531

```

## (2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 274 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: SAC\_23836

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

GGGGGATCCG CCCGATCATG ATGAATCCCC TGGGACCAAG CTATAACACA CAGTCCTCTG	60
AGAAAAAGGA TTGTGGTTTG CAGCAAATTC AGCCAATTTT GTCAAGCAAT TTGTGAGAGA	120
CACTCCTTAC TTGGTTGTTT CCTTCTCAGT CTTACGGGGT GCCTTATCAC TCCGTGATGG	180
TTCATGGCCA CACAGTCATA TTTCAGGGAG AAGTCCTTGT CGGTCACACC AGTTATCCTT	240
AACAGTGAGG TTAAGCAAAT CATGCCATTG CTGG	274

(2) INFORMATION FOR SEQ ID NO:97:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 254 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_23853

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

GGGGGATCCG CCCGATCGTG TTATTGAGGT CATCCACGGT GGAGATTAGC AGGTGACTCC	60
GTTTTCTCTG AAATGAGTGG AGTTGGGAAG GGGACCCGTG GCCGTGCTCC AGCTATAGTT	120
GGTGGGCGGT GGTTCGCTGC GAGCTTCACA GGTCAGGTTC ACGTTAGTAG GCCCAATGTA	180
CCAGTTGCCT TCATAGCCAG AGATGGACAC TTCGGGTGGA TAAGGTAGGG AAAGGATCGG	240
GCTAGAGCGG CCGC	254

(2) INFORMATION FOR SEQ ID NO:98:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 255 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_23906

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GGGGGATCCG CCCACAGAAC CCAGAGGAAG GAGAGGCTGC TGGGGTGGAG GCCTAGGCGC	60
TGGAGACATG TGGAGTTCTC TAGGGGTCTG CAGCAACCTC GGAAAGCTGG GAGATTCCCT	120
CCTTGAGACT CCTACATATA GAAACTGAT GCTTCTGTCT CATTCCATGC GGCTTTTCCT	180
GCGGTATTCC TGTAGCGCTT TCTCTGCCAC TGTGTCCATA AACTTAGGGT TATCCTTGGA	240
GACTTCTTCT GGTAC	255

## (2) INFORMATION FOR SEQ ID NO:99:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 219 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vii) IMMEDIATE SOURCE:  
 (B) CLONE: SAC\_23907

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

GGGGGATCCG CCCACCCCAT CTAATCCTCT GAGGCTCAGG GTGACTCATT CTTCTTGGA	60
CCCAGGGGCG GGGCCGACAT TTGTTCCCCC AGTTGGGCCC TGCTGCCCCA GGCCAGTGGC	120
ATTCCCAGTG TCTCCAGCCA CTAAAGCCAC ATTCTCAGG TAGTTGGTAT TGAAGCAGTT	180
GGCTTGTTTA TCTCCAGGAG ACAAGGTTTCG AGCGGCCGC	219

## (2) INFORMATION FOR SEQ ID NO:100:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 182 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

- (vii) IMMEDIATE SOURCE:  
 (B) CLONE: SAC\_23916

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

GGGGGATCCG CCCTCGAGCG GCCGCCACA AGAGATTGAA CCTAGAGCTG TGTGCCTCCT	60
--	----



GGGCAAGCAC TCTCTACTGA GCTACATCCC TGTTAAAGTG CCTTTTTTGG GAGCTTTGTC 120  
TTCCAGCCTG CCAATCAACC CACTTTATGG GTGTGCCTAG ATTCCCCTTT CTCTGAGTAG 180  
GG 182

(2) INFORMATION FOR SEQ ID NO:101:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 191 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_23922

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

GGGGGATCCG CCCACTGTTC TCAAACAGTT GCCTAACAAT ACATTTAATT AAGAAATTAT 60  
CATAACTGTA TGCTTACACA TATGTGGATA GGCATCAATG AAGATTACAA CAGCTCAGCC 120  
ACAGTGCTGC AATGCTCTAC ACTACTACAA AACACCCACC ACGGGCCTGN NACCTGCCCT 180  
CGAGCGGCCG C 191

(2) INFORMATION FOR SEQ ID NO:102:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 234 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_24030

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

GGGGGATCCG CCCACCATAT CTATACGGCC CTGGAAGAGG AGATAGAGCG AAACAAGCAG 60  
AACCCAGTCT ATGCCCCGCT CTACTTNCCT GAGGAGCTGC ACCGAAGGCT GCCCTAGAGC 120  
AGGACATGGC CTTCTGGTAT GGGCCCCACT GGCAGGAGGC CATCCCTTAN NNACCAGCCA 180  
CACAGCACTA CGTAAAGCGT CTCCACGAGG TGGGAGGTGG GCTAGAGCGG CCGC 234

(2) INFORMATION FOR SEQ ID NO:103:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 172 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
    (B) CLONE: SAC\_24078

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GGGGGATCCG CCCGATCAAG ATGACTAAGA TGCTCAAAGG ATTCGACGCT GTGGGAAATG	60
CCACAGGTTT CCGGTTTCGCC TACACCCAG CCATGAGAGC CTCTGTGGAT ATGTCCACAA	120
GTCCCAGAAC CCGCAGCGAG GAGTTTCTCA TCGCGGGCCG TTTAAGGAAC GG	172

(2) INFORMATION FOR SEQ ID NO:104:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 241 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
    (B) CLONE: SAC\_24105

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

GGGGGATCCG CCCGATCTGT GCATGTGAGC CGAACAGAAG TCAGCAATAA CCATGTCTTG	60
ATTTACCTGG ATAAGGTGTC AAATCAGACG GTGAACTTGT CCTTCACGGC TCAGCAAGAT	120
ATTCCAATAA GGGACCTGAA GCCAGCCGTA GTGAAAGTCT ACGATTACTA TGAGAAAGAT	180
GAGTTTGCAG TTGCAAATA CAGCGCTCCC TGCAGCACAG ATTATGGAAA TGCCTGAGGA	240
C	241

(2) INFORMATION FOR SEQ ID NO:105:

- (i) SEQUENCE CHARACTERISTICS:  
    (A) LENGTH: 234 base pairs  
    (B) TYPE: nucleic acid  
    (C) STRANDEDNESS: double  
    (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_24140

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

GGGGGATCCG CCCGATCCTT AAAGGTCCCA TAACCAAGTA CTCCATGCAG TCATCTCCTC	60
TAACCTGTGG GTCCCCTCCA TTCATTGCAA GCTGCTGGAC ATAGCCTGCT GGGTCCACCA	120
CAAGTACAAC AGTGACAAGT CCAGCACCTA TGTGAAGAAT GGCACATCCT TCGACATCCA	180
CTACGGCTCA GGTAGCCTCT CTGGGTACCT GAGCCAGGAC ACTGTGTTGG TTCC	234

(2) INFORMATION FOR SEQ ID NO:106:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 217 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_24142

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

GGGGGATCCG CCCGATCTAT CCGGGTATAT TCGGAAGTGT GGAGTCTGAC CGCAGTGCAC	60
TGCTACTCAG CTCACTCTGT CTCTTTCTCA CTGTGGGAGT AGCTATGGGA AAGGTCACAC	120
TTCCTTGCTT TCCCCTCAAA CCAACAGGGC TTTCACCTCT CAGTCTCCCA ACGCACTGCA	180
CACTGCTCCC AGCATCCCGC ACAGTCTCAC CACACAC	217

(2) INFORMATION FOR SEQ ID NO:107:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 172 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO

- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_24192

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

GGGGGATCCG CCCACATCTC AGCGGTGGGG ACTCAGACAT TCCTGTGTTC CCTCCTTGGC 60  
CTCCAGCCTC TCTGTAGGAA CCTCCAGCAG CCTGCCACCA GATTTCCTT AGCTTCCACT 120  
GTCTCCATGA GCTTTAAAT GTGGGCGGCC GCTCGAGGGC TAGAGCGGCC GC 172

## (2) INFORMATION FOR SEQ ID NO:108:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 212 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_24296

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

GGGGGATCCG CCCACTTGGC TGACCGCCAG GCGAGTGACT GTGACTCCAG GTTCGCCCAC 60  
TACCGCAGCA GATACTGCCT CGCCTATGCC ACCTTCATAG TAGTGGTCCT CCACGGTGAG 120  
GATCCTGCCT TTGGTTGCTC TGGCACAGTC GAGAATGAGC TTTTGTCCA GGGGCTTGAT 180  
GGTGAAGGGG TCCAGTGGGC TAGAGCGGCC GC 212

## (2) INFORMATION FOR SEQ ID NO:109:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 210 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:  
(B) CLONE: SAC\_24445

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

GGGGGATCCG CCCACAGCAT TCGGGTATGA GGCAAGGGGA GTGTTCATTC ACACACACAG 60  
TAGCTTCGGG TCTGTGAGGT CTCCTTGTTT CCGGGTAGGT TTCTAAAGAC GGAGAAAAAA 120

ATGATTCTGG TTATCAAGAC TACTGTGACC GTATTAGATC CCAGAATGGG CAAGCATCAG 180  
TGTGTGACCA TCCGAACAAA AGGAATTTTT 210

(2) INFORMATION FOR SEQ ID NO:110:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 200 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_24527

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

GGGGGATCCG CCCACCATGG TCTCACTCAG AGACGGGTTT GGTCAGTTTT AAGATGAACT .60  
CCTTCGGATG TAGTGTCCTG AAGTCCTTAT TATCGCTGTC CATCAATTCA CGATAGGATA 120  
GCTGGCTGCG GTGGCAAGTC CGCAGTGGTT GTTCCACCTG CCCACATTAT TCTCTCGTGG 180  
CCATGGTTCT TCCCTTAGAG 200

(2) INFORMATION FOR SEQ ID NO:111:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 400 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (vii) IMMEDIATE SOURCE:
  - (B) CLONE: SAC\_24540

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

GGGGGATCCG CCCACAGTGG GAAGGATGGA AATAGCTGTC ATCTCTCAA AACTGAGTTC 60  
CTTTCCTTCA TGAACACGGA GCTGGCCGCC TTCACGAAGA ACCAGAAGGA CCCCGGTGTC 120  
CTCGACCGCA TGATGAAGAA GCTGGACCTC AACAGTGATG GGCAGCTAGA TTTCCAAGAG 180  
TTTCTCAACC TTATTGGTGG CTTAGCTATA GCACCTGCCC ACTCTCGTGG CTGGCCCTCT 240  
GCTTATACTC CTTGGACTGC TCCAGGCTGT GTGTTTCCAC GCTTGACCTG CCCACAGAGC 300  
CAAAAGAGAC ACATTGGCTA CTTTAACCAC CTAAAGCGG ACTCCAGGAA TATCACCCAC 360

GGCATGGCCT TTTGCGAACA AATCCCANCA ACCAAAAAAA

400

## (2) INFORMATION FOR SEQ ID NO:112:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vii) IMMEDIATE SOURCE:

(B) CLONE: SAC\_24623

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

GGGGGATCCG CCCACAGTGC ATATTGGCGG CGCTCGCCTC ATTACGATTC GCCAGCTTGC	60
TTCTCTTGTT CAATTGTTTC TTTTGAAGGC AGAGGATTTT TCTCTGTGT TTCTGTCTTC	120
TTCAACTTCG ACTTATCGAA TTTCACCTGC CCACTAAACC CACAGCCCAG CCCTCCTTCC	180
TTCTCCTGT CCTGGAATTG	200

What is claimed is:

1. Purified and isolated nucleic acid encoding all or a unique fragment of a KIM.
2. Nucleic acid of claim 1 that is RNA.
3. Nucleic acid of claim 1 that is DNA.
- 5 4. Nucleic acid of claim 3 that is cDNA.
5. Nucleic acid that is complementary to the nucleic acid of claim 1.
6. Nucleic acid that hybridizes to the nucleic acid of claim 1 or 5.
7. Nucleic acid of claim 1 that is in operative association with an expression control element therefor.
- 10 8. Antisense nucleic acid of claim 1 that is sufficient, when internalized within a cell, to disrupt expression of a cellular KIM gene.
9. Purified and isolated nucleic acid encoding all or a unique fragment of a KIM, said nucleic acid having a sequence selected from the sequences listed in TABLE 1.
10. Purified and isolated nucleic acid encoding all or a unique fragment of a KIM, said  
15 nucleic acid having a sequence selected from degenerate variants of the sequences listed in TABLE 1.
11. Purified and isolated nucleic acid encoding all or a unique fragment of a KIM, said nucleic acid having a sequence selected from homologous variants of the sequences listed in TABLE 1.
- 20 12. Purified and isolated nucleic acid encoding a KIM fusion polypeptide, the sequence of said nucleic acid comprising a KIM sequence selected from the sequences listed in TABLE 1 fused in frame to a sequence encoding a non- KIM polypeptide.
13. Nucleic acid of claim 12 wherein the encoded non-KIM polypeptide is selected  
25 from: a secretable leader polypeptide, an immunoglobulin polypeptide, a binding pair partner, a toxin or toxoid, an enzyme, and a detectable polypeptide.

14. A vector comprising a KIM-encoding nucleic acid of any one of claims 1 to 13.
15. A host cell stably transformed or transfected with a vector of claim 14.
16. A method for producing a KIM polypeptide, comprising the steps of growing a host cell of claim 15 under conditions sufficient for the production of polypeptides of vector origin; and, recovering an expressed KIM polypeptide.
17. A purified and isolated polypeptide comprising all or a unique fragment of a KIM, said polypeptide having a sequence selected from the sequences encoded by nucleic acids listed in TABLE 1.
18. A purified and isolated polypeptide comprising all or a unique fragment of a KIM, said polypeptide having a sequence selected from variants of the sequences encoded by nucleic acids listed in TABLE 1.
19. A polypeptide of claim 18 wherein said variant is a splice variant, truncation variant, or substitution variant of an encoded TABLE 1 sequence.
20. A KIM fusion polypeptide, the sequence of said polypeptide comprising all or a unique fragment of an encoded KIM sequence listed in TABLE 1, fused to a non-KIM polypeptide sequence.
21. A fusion polypeptide of claim 20 wherein the non-KIM polypeptide sequence is selected from: a secretable leader polypeptide sequence, an immunoglobulin polypeptide sequence, a binding pair partner sequence, a toxin or toxoid sequence, an enzyme sequence, and a detectable polypeptide sequence.
22. A conjugate comprising a KIM polypeptide of claim 17, 18 or 19, linked to a detectable moiety, imageable moiety, radiolabeled moiety, or toxin moiety.
23. A KIM fusion polypeptide, the sequence of said polypeptide comprising all or a unique fragment of an encoded KIM sequence selected from nucleic acid sequences listed in TABLE 2, fused to a non-KIM polypeptide sequence.
24. A fusion polypeptide of claim 23 wherein the non-KIM polypeptide sequence is selected from: a secretable leader polypeptide sequence, an immunoglobulin



polypeptide sequence, a binding pair partner sequence, a toxin or toxoid sequence, an enzyme sequence, and a detectable polypeptide sequence.

25. A conjugate comprising a KIM polypeptide of claim 23 or 24, linked to a detectable moiety, imageable moiety, radiolabeled moiety, or toxin moiety.
- 5 26. An antibody that binds selectively to all or a unique fragment of a KIM polypeptide encoded by a nucleic acid sequence listed in TABLE 1.
27. A conjugate comprising an antibody of claim 26, linked to a detectable moiety, imageable moiety, radiolabeled moiety, or toxin moiety.
28. An antibody that binds selectively to all or a unique fragment of a KIM polypeptide  
10 encoded by a nucleic acid sequence listed in TABLE 2.
29. A conjugate comprising an antibody of claim 29, linked to a detectable moiety, imageable moiety, radiolabeled moiety, or toxin moiety.
30. Use in therapy of an antisense KIM nucleic acid of claim 8.
31. Use in therapy of a KIM vector of claim 14.
- 15 32. Use in therapy of a KIM polypeptide of claim 17, 18 or 19.
33. Use in therapy of a KIM fusion protein of claim 20 or 21, or of a KIM conjugate of claim 22.
34. Use in therapy of an antibody of claim 26, or of an antibody conjugate of claim 27.
35. A pharmaceutical composition comprising a physiologically acceptable carrier  
20 having dispersed therein, to a therapeutically effective concentration: a KIM nucleic acid of claim 8; a KIM vector of claim 14; a KIM polypeptide of claim 17, 18 or 19; a KIM fusion protein of claim 20 or 21; a KIM conjugate of claim 22; an anti-KIM antibody of claim 26; or, an anti-KIM antibody conjugate of claim 27.
36. A method for detecting the presence or status of renal injury or of renal disease,  
25 comprising the step of measuring the concentration of a KIM polypeptide in a

sample selected from serum, urine or urine sediment of an individual afflicted with, or at risk of developing, renal injury or an impairment of renal function.

37. A method for detecting the presence or status of an abnormal response to tissue injury or other stimulus, comprising the step of measuring the concentration of a KIM polypeptide in a sample selected from serum, urine or urine sediment of an individual afflicted with, or at risk of developing, an autoimmune response or abnormal tissue growth arising from or affecting renal tissue.
38. A method according to claim 36 or 37, wherein said measuring step involves contacting said sample with an anti-KIM antibody of claim 26 or 28, or with a conjugate of claim 27 or 29.
39. A kit for practice of the method of claim 36 or 37, comprising at least one reagent selected from: the anti-KIM antibody of claim 26; the anti-KIM antibody of claim 28; the conjugate of claim 27; and, the conjugate of claim 29.
40. A method for detecting the presence or status of renal injury or of renal disease, comprising the step of measuring the level of a KIM nucleic acid in a sample selected from serum, urine or urine sediment of an individual afflicted with, or at risk of developing, renal injury or an impairment of renal function.
41. A method for detecting the presence or status of an abnormal response to tissue injury or other stimulus, comprising the step of measuring the level of a KIM nucleic acid in a sample selected from serum, urine or urine sediment of an individual afflicted with, or at risk of developing, an autoimmune response or abnormal tissue growth arising from or affecting renal tissue.
42. A method according to claim 40 or 41, wherein said measuring step involves contacting said sample with a KIM nucleic acid of claim 9, 10 or 11; or with a nucleic acid probe that hybridizes to KIM nucleic acid of claim 9, 10 or 11; or with a nucleic acid having a sequence selected from sequences listed in TABLE 2; or with a nucleic acid probe that hybridizes to a nucleic acid listed in TABLE 2.

43. A kit for practice of the method of claim 40 or 41, comprising at least one reagent selected from: the KIM nucleic acid of claim 9, 10 or 11; or a nucleic acid probe that hybridizes to KIM nucleic acid of claim 9, 10 or 11; or a nucleic acid having a sequence selected from sequences listed in TABLE 2; or a nucleic acid probe that hybridizes to a nucleic acid listed in TABLE 2.
44. A method of imaging cells or tissue expressing or producing a KIM, comprising the steps of contacting the cells or tissue with an imageable KIM-binding reagent; and, imaging an accumulation of the KIM-binding reagent.
45. A method according to claim 44 wherein the cells or tissue are disposed in body tissue of an individual suspected of harboring an imageable locus of KIM expression and/or production.
46. A method of locating a tissue mass abnormally producing or expressing a KIM in an individual suspected of affliction with an abnormal tissue response to injury or other stimulus, comprising the step of: administering an imageable KIM-binding reagent to the individual; and, imaging said tissue mass within the body of the individual.
47. A method according to claim 44 wherein the tissue mass is a tumor arising from or affecting renal tissue of the individual.
48. A method of treating an individual afflicted with, or at risk of developing, a disease or condition contributed to or associated with a dysfunction or dysregulation of a KIM gene or protein, comprising the step of administering to the individual a pharmaceutical composition of claim 35.
49. A method of treating an individual afflicted with, or at risk of developing, a disease or condition beneficially affected by therapeutic administration of a KIM composition of the present invention, comprising the step of administering to the individual a pharmaceutical composition of claim 35.

50. A method according to claim 48 or 49 wherein the disease is a renal disease; or, the condition is an impairment of renal function.
51. A method of inhibiting the growth of KIM-expressing tumor cells, comprising the step of contacting the tumor cells with a pharmaceutical composition comprising  
5 involving the step of contacting the cells with: an antisense KIM nucleic acid of claim 8; a KIM fusion protein of claim 20, 21, 23 or 24; a KIM conjugate of claim 22 or 25; an antibody of claim 26 or 28; or an antibody conjugate of claim 27 or 29.
52. A method according to claim 36, 37, 40, 41, 45, 46, 47, 48, 49, or 50 wherein the individual is a human.

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/10547

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/62 C12N5/10 C12N1/21 C07K14/47  
C07K16/18 C12Q1/68 A61K31/70 A61K38/17 A61K39/395  
A61K48/00 G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C07K C12Q A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	N.H. LEE ET AL.: "Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment" EMBL SEQUENCE DATABASE, 1 October 1995, XP002080245 Heidelberg, FRG EST109574 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA clone PRNAT01 3' end, Accession no. H33500; --- -/-	5,6



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

### \* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

12 October 1998

Date of mailing of the international search report

22/10/1998

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	N.H. LEE ET AL.: "Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment" EMBL SEQUENCE DATABASE, 1 October 1995, XP002080246 Heidelberg, FRG EST111059 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA clone PRNBQ59 5' end; Accession no. H34282; ---	6
X	N.H. LEE ET AL.: "Comparative expressed-sequence-tag analysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment" EMBL SEQUENCE DATABASE, 1 October 1995, XP002080247 Heidelberg, FRG EST111810 Rat PC-12 cells, NGF-treated (9 days) Rattus sp. cDNA clone RPNCM27 5' end, Accession no. H34674; ---	6
A	US 5 552 313 A (CALVET JAMES P ET AL) 3 September 1996 see the whole document ---	1-52
A	WO 95 26732 A (AMUR RESEARCH CORP ; CHASLOW FRED I (US)) 12 October 1995 see the whole document ---	1-52
A	BONALDO FATIMA DE M ET AL: "NORMALIZATION AND SUBTRACTION: TWO APPROACHES TO FACILITATE GENE DISCOVERY" GENOME RESEARCH, vol. 6, no. 9, September 1996, pages 791-806, XP002039972 see the whole document ---	1-52
P,X	WO 97 44460 A (BIOGEN INC ; SANICOLA NADEL MICHELE (US); BONVENTRE JOSEPH V (US);) 27 November 1997 cited in the application see the whole document ---	1-8, 35-52
P,X	M. TAKADA ET AL.: "The cytokine-adhesion molecule cascade in ischemia/perfusion injury of the rat kidney. Inhibition by a soluble P-selectin ligand" J. CLIN. INVEST., vol. 99, no. 11, 1 June 1997, pages 2682-2690, XP002079925 ROCKEFELLER UNIVERSITY PRESS, NEW YORK, US see the whole document ---	1-8

-/--

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 98/10547

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	<p>T. ICHIMURA ET AL.: "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion molecule containing a novel immunoglobulin domain, is up-regulated in renal cells after injury"</p> <p>J. BIOL. CHEM., vol. 273, no. 7, 13 February 1998, pages 4135-4142, XP002079926</p> <p>AM. SOC. BIOCHEM. MOL.BIOL., INC., BALTIMORE, US cited in the application see the whole document</p> <p>-----</p>	1-8

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 98/10547

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
Please see Further Information sheet enclosed.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.



Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 5552313	A	03-09-1996	NONE	
WO 9526732	A	12-10-1995	AU 2206495 A	23-10-1995
WO 9744460	A	27-11-1997	AU 3567697 A	09-12-1997

1. *Chlorophyll*

1 2

1.

100